Refine Search

Search Results -

Terms	Documents
701/209	2256

Database:
US Pre-Grant Publication Full-Text Database
US Patents Full-Text Database
US OCR Full-Text Database
EPO Abstracts Database
JPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

Search: L28

Refine Search

Recall Text Clear Interrupt

Search History

DATE: Saturday, January 20, 2007 **Purge Queries** Printable Copy Create Case

Set Name side by side	Query	Hit Count	Set Name result set
	DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI,TDBD; PLUR=YES; OP=OR		
L28	701/209	2256	L28
L27	701/200	3355	L27
L26	701.clas.	39615	L26
L25	707/200	5559	L25
L24	707/1	9007	L24
L23	707/104.1	7599	L23
L22	707.clas.	40545	L22
L21	340/995	2904	L21
L20	340/991	753	L20
L19	340/990	2477	L19
L18	340.clas.	163055	L18
	DB=USPT; PLUR=YES; OP=OR		
L17	'5617319'.pn.	1	L17

L16	'4962458'.pn.	1	L16
L15	'5815161'.pn.	1	L15
L14	'5815161'.pn.	1	L14
L13	'5832406'.pn.	1	L13
DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI,TDBD; PLUR=YES; OP=OR			
L12	l6 and (kd-tree or "peano tree")	3	L12
DB=USPT; PLUR=YES; OP=OR			
L11	("5802492"| "5559707")[URPN]	258	L11
L10	(5231584 | 5559707 | 5377113 | 4939662 | 5270937 | 5353034 | 5543789 | 5208756 | 4926336)![PN]	9	L10
L9	("5802492"| "5559707")[PN]	2	L9
DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI,TDBD; PLUR=YES; OP=OR			
L8	L7 and (index or indices)	6	L8
L7	L6 and rectangles	6	L7
L6	L5 and parcels	14	L6
L5	L4 and geographic same features	74	L5
L4	L3 and map same (database or data with base)	230	L4
L3	geographical near (database or data with base)	527	L3
DB=EPAB,JPAB; PLUR=YES; OP=OR			
L2	L1 and map same (database or data with base)	7	L2
L1	geographical near (database or data with base)	35	L1

END OF SEARCH HISTORY

Refine Search

Search Results -

Terms	Documents
L3 not @py>1998	34

Database:
US Pre-Grant Publication Full-Text Database
US Patents Full-Text Database
US OCR Full-Text Database
EPO Abstracts Database
JPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

Search:

Refine Search

Recall Text Clear Interrupt

Search History

DATE: Saturday, January 20, 2007 Purge Queries Printable Copy Create Case

Set Name side by side	Query	Hit Count	Set Name result set
	DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI,TDBD; PLUR=YES; OP=OR		
L10	l3 not @py>1998	34	L10
	DB=USPT; PLUR=YES; OP=OR		
L9	'5710978'.pn.	1	L9
L8	'5719563'.pn.	1	L8
L7	'5734696'.pn.	1	L7
L6	'5768689'.pn.	1	L6
	DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI,TDBD; PLUR=YES; OP=OR		
L5	l4 and (index or indices)	5	L5
L4	L3 and sub-areas	9	L4
L3	L2 and parcels	337	L3
L2	L1 and (map or atlas)	7141	L2
L1	(geographic or geographical) same (database or data with base)	15746	L1

END OF SEARCH HISTORY

IEEE Xplore® RELEASE 2.1

Home | Login | Logout | Access Information | Alerts |

Welcome United States Patent and Trademark Office

Search Results

BROWSE SEARCH IEEE XPLORE GUIDE

Results for "((database<in>metadata) <and> (kd-tree<in>metadata))"

e-mail

Your search matched **11** of **1472243** documents.

A maximum of **100** results are displayed, **25** to a page, sorted by **Relevance** in **Descending** order.

» Search Options

View Session History

New Search

» Key

IEEE JNL	IEEE Journal or Magazine
IEE JNL	IEE Journal or Magazine
IEEE CNF	IEEE Conference Proceeding
IEE CNF	IEE Conference Proceeding
IEEE STD	IEEE Standard

Modify Search

((database<in>metadata) <and> (kd-tree<in>metadata)) Search

Check to search only within this results set

Display Format: Citation Citation & Abstract

view selected items Select All Deselect All

1. **Ω-storage: a self organizing multi-attribute storage technique for very lar memories**
Karlsson, J.S.; Kersten, M.L.;
Database Conference, 2000. ADC 2000. Proceedings. 11th Australasian
31 Jan.-3 Feb. 2000 Page(s):57 - 64
Digital Object Identifier 10.1109/ADC.2000.819814
AbstractPlus | Full Text: PDF(100 KB) **IEEE CNF**
Rights and Permissions

2. **Increasing retrieval efficiency by index tree adaptation**
Tagare, H.D.;
Content-Based Access of Image and Video Libraries, 1997. Proceedings. IEEE
20 June 1997 Page(s):28 - 35
Digital Object Identifier 10.1109/IVL.1997.629717
AbstractPlus | Full Text: PDF(600 KB) **IEEE CNF**
Rights and Permissions

3. **Using triangle inequality to efficiently process continuous queries on hig streaming time series**
Zhengrong Yao; Like Gao; Wang, X.S.;
Conference on Scientific and Statistical Database Management, 2003. 15th Int
9-11 July 2003 Page(s):233 - 236
Digital Object Identifier 10.1109/SSDM.2003.1214985
AbstractPlus | Full Text: PDF(266 KB) **IEEE CNF**
Rights and Permissions

4. **Content based image retrieval: related issues using Euler vector**
Bishnu, A.; Bhunre, P.K.; Bhattacharya, B.B.; Kundu, M.K.; Murthy, C.A.; Acha
Image Processing. 2002. Proceedings. 2002 International Conference on
Volume 2, 22-25 Sept. 2002 Page(s):II-585 - II-588 vol.2
Digital Object Identifier 10.1109/ICIP.2002.1040018
AbstractPlus | Full Text: PDF(480 KB) **IEEE CNF**
Rights and Permissions

5. **Retrieve of images by content**
Djeraba, C.; Savory, I.; Briand, H.;
Intelligence and Systems, 1996., IEEE International Joint Symposia on
4-5 Nov. 1996 Page(s):261 - 267

Digital Object Identifier 10.1109/IJSIS.1996.565077

AbstractPlus | Full Text: PDF(924 KB) **IEEE CNF**
Rights and Permissions

☐ 6. **Dynamic query range for multikey searching**
Sun, X.-H.; Kamel, N.N.;
Computer Software and Applications Conference, 1990. COMPSAC 90. Proce
Fourteenth Annual International
31 Oct.-2 Nov. 1990 Page(s):118 - 123
Digital Object Identifier 10.1109/CMPSAC.1990.139338

AbstractPlus | Full Text: PDF(516 KB) **IEEE CNF**
Rights and Permissions

☐ 7. **Euler vector for search and retrieval of gray-tone images**
Bishnu, A.; Bhattacharya, B.B.; Kundu, M.K.; Murthy, C.A.; Acharya, T.;
Systems, Man and Cybernetics, Part B, IEEE Transactions on
Volume 35, Issue 4, Aug. 2005 Page(s):801 - 812
Digital Object Identifier 10.1109/TSMCB.2005.846642

AbstractPlus | Full Text: PDF(1336 KB) **IEEE JNL**
Rights and Permissions

☐ 8. **Efficient Indexing, Color Descriptors and Browsing in Image Databases**
Batalas, N.; Diou, C.; Delopoulos, A.;
Semantic Media Adaptation and Personalization, 2006. SMAP '06. First Interna
on
Dec. 2006 Page(s):129 - 134
Digital Object Identifier 10.1109/SMAP.2006.15

AbstractPlus | Full Text: PDF(464 KB) **IEEE CNF**
Rights and Permissions

☐ 9. **Using bitmap index for interactive exploration of large datasets**
Kesheng Wu; Koegler, W.; Chen, J.; Shoshani, A.;
Conference on Scientific and Statistical Database Management, 2003. 15th Int
9-11 July 2003 Page(s):65 - 74
Digital Object Identifier 10.1109/SSDM.2003.1214955

AbstractPlus | Full Text: PDF(2915 KB) **IEEE CNF**
Rights and Permissions

☐ 10. **Multiattribute indexing with m-Q-trees**
Polo, A.; Mansanzo, J.C.; Salas, M.; Sanchez, M.; Jurado, E.; Barrena, M.; He
Martinez, J.M.;
Parallel and Distributed Processing, 1998. PDP '98. Proceedings of the Sixth E
Workshop on
21-23 Jan. 1998 Page(s):94 - 101
Digital Object Identifier 10.1109/EMPDP.1998.647185

AbstractPlus | Full Text: PDF(776 KB) **IEEE CNF**
Rights and Permissions

☐ 11. **Join strategies on KD-tree indexed relations**
Kitsuregawa, M.; Harada, L.; Takagi, M.;
Data Engineering, 1989. Proceedings. Fifth International Conference on
6-10 Feb. 1989 Page(s):85 - 93
Digital Object Identifier 10.1109/ICDE.1989.47203

AbstractPlus | Full Text: PDF(700 KB) **IEEE CNF**
Rights and Permissions

IEEE Xplore RELEASE 2.1

Home | Login | Logout | Access Information | Alerts |

Welcome United States Patent and Trademark Office

Search Session History

BROWSE **SEARCH** **IEEE XPLORE GUIDE**

Sat, 20 Jan 2007, 10:13:39 AM EST

Edit an existing query or compose a new query in the Search Query Display.

Select a search number (#) to:

- Add a query to the Search Query Display
- Combine search queries using AND, OR, or NOT
- Delete a search
- Run a search

Search Query Display

Recent Search Queries

#1	((geographic database<in>metadata) <and> (map<in>metadata))
#2	((geographic database<in>metadata) <and> (map<in>metadata))
#3	((geographic database<in>metadata) <and> (map<in>metadata))
#4	((geographic database<in>metadata) <and> (map<in>metadata))
#5	((geographic database<in>metadata) <and> (map<in>metadata))
#6	((geographic database<in>metadata) <and> (map<in>metadata))
#7	((geographic database<in>metadata) <and> (kd-tree<in>metadata))
#8	((database<in>metadata) <and> (kd-tree<in>metadata))
#9	((database<in>metadata) <and> (kd-tree<in>metadata))
#10	((database<in>metadata) <and> (kd-tree<in>metadata))
#11	((database<in>metadata) <and> (kd-tree<in>metadata))

Indexed by Inspec

Help Contact Us Privacy & :

Results for **"((geographic database<in>metadata) <and> (map<in>metadata))"**

e-mail

Your search matched **14** of **1472243** documents.

A maximum of **100** results are displayed, **25** to a page, sorted by **Relevance** in **Descending** order.

» Search Options

View Session History

New Search

» Key

IEEE JNL	IEEE Journal or Magazine
IEE JNL	IEE Journal or Magazine
IEEE CNF	IEEE Conference Proceeding
IEE CNF	IEE Conference Proceeding
IEEE STD	IEEE Standard

Modify Search

((geographic database<in>metadata) <and> (map<in>metadata)) Search

☐ Check to search only within this results set

Display Format: ◉ Citation ○ Citation & Abstract

view selected items Select All Deselect All

☐ 1. **Use of geographic information systems and photogrammetric techniques NC division of forest resources pre-suppression fire planning and forest**
Pickens, L.M.; Cheshire, H.M.; Devine, H.A.;
Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2(International
Volume 6, 24-28 July 2000 Page(s):2712 - 2714 vol.6
Digital Object Identifier 10.1109/IGARSS.2000.859690
AbstractPlus | Full Text: PDF(276 KB) **IEEE CNF**
Rights and Permissions

☐ 2. **Unsupervised connectionist clustering algorithms for a better supervisec application to a radio communication problem**
Bougrain, L.; Alexandre, F.;
Neural Networks, 1999. IJCNN '99. International Joint Conference on
Volume 5, 10-16 July 1999 Page(s):3451 - 3456 vol.5
Digital Object Identifier 10.1109/IJCNN.1999.836220
AbstractPlus | Full Text: PDF(416 KB) **IEEE CNF**
Rights and Permissions

☐ 3. **A road extraction method from topographical map images**
Chang Ahn; Kyu Sik Kim; Sang Burm Rhee; Kye Young Lee;
Communications, Computers and Signal Processing, 1997. '10 Years PACRIM Networking the Pacific Rim'. 1997 IEEE Pacific Rim Conference on
Volume 2, 20-22 Aug. 1997 Page(s):839 - 842 vol.2
Digital Object Identifier 10.1109/PACRIM.1997.620390
AbstractPlus | Full Text: PDF(516 KB) **IEEE CNF**
Rights and Permissions

☐ 4. **Browsing in geographic databases: an object-oriented approach**
Clementini, E.; D'Atri, A.; Di Felice, P.;
Visual Languages, 1990., Proceedings of the 1990 IEEE Workshop on
4-6 Oct. 1990 Page(s):125 - 131
Digital Object Identifier 10.1109/WVL.1990.128393
AbstractPlus | Full Text: PDF(496 KB) **IEEE CNF**
Rights and Permissions

☐ 5. **An object-oriented conceptual model for the representation of geographi**
Clementini, E.; Di Felice, P.; D'Atri, A.;

Applied Computing, 1991., [Proceedings of the 1991] Symposium on
3-5 April 1991 Page(s):472 - 480
Digital Object Identifier 10.1109/SOAC.1991.143923

AbstractPlus | Full Text: PDF(660 KB) **IEEE CNF**
Rights and Permissions

6. **Personal dynamic maps based on distributed geographic information sei**
Arikawa, M.;
Vehicle Navigation and Information Systems Conference, 1994. Proceedings.,
31 Aug.-2 Sept. 1994 Page(s):591 - 596
Digital Object Identifier 10.1109/VNIS.1994.396784

AbstractPlus | Full Text: PDF(464 KB) **IEEE CNF**
Rights and Permissions

7. **An Indexing System for Photos Based on Shooting Position and Orientat Geographic Database**
Iwasaki, K.; Yamazawa, K.; Yokoya, N.;
Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on
6-8 July 2005 Page(s):390 - 393
Digital Object Identifier 10.1109/ICME.2005.1521442

AbstractPlus | Full Text: PDF(3248 KB) **IEEE CNF**
Rights and Permissions

8. **Dynamic W-CDMA network planning using mobile location**
Sharma, S.; Nix, A.R.;
Vehicular Technology Conference, 2002. Proceedings. VTC 2002-Fall. 2002 IE
Volume 2, 24-28 Sept. 2002 Page(s):1182 - 1186 vol.2
Digital Object Identifier 10.1109/VETECF.2002.1040791

AbstractPlus | Full Text: PDF(346 KB) **IEEE CNF**
Rights and Permissions

9. **Cooperation and fusion of operators: Application to automatic matching objects**
Dherete, P.; Desachy, J.;
Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2(International
Volume 6, 24-28 July 2000 Page(s):2626 - 2628 vol.6
Digital Object Identifier 10.1109/IGARSS.2000.859662

AbstractPlus | Full Text: PDF(320 KB) **IEEE CNF**
Rights and Permissions

10. **WAVEMAP: A geographic information system for wave modelling in coas**
Opishinski, T.B.;
OCEANS '94. 'Oceans Engineering for Today's Technology and Tomorrow's P Proceedings
Volume 1, 13-16 Sept. 1994 Page(s):I/872 - I/877 vol.1
Digital Object Identifier 10.1109/OCEANS.1994.363976

AbstractPlus | Full Text: PDF(496 KB) **IEEE CNF**
Rights and Permissions

11. **Environmental-geological study of coastal areas of Salento Peninsula (S using multitemporal aerial photography and satellite imagery**
Ricchetti, E.; Ricchetti, G.; Sani, R.;
Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedinc International
Volume 6, 2004 Page(s):3642 - 3645 vol.6
Digital Object Identifier 10.1109/IGARSS.2004.1369908

AbstractPlus | Full Text: PDF(1235 KB) **IEEE CNF**
Rights and Permissions

12. **Cadastre map assembling: a puzzle game resolution**
Viglino, J.-M.; Guigues, L.;
Document Analysis and Recognition, 2001. Proceedings. Sixth International C
10-13 Sept. 2001 Page(s):1235 - 1239
Digital Object Identifier 10.1109/ICDAR.2001.953979
AbstractPlus | Full Text: PDF(608 KB) **IEEE CNF**
Rights and Permissions

13. **Modelling the impact of land use planning and management practices on fragmentation of urban landscapes**
McCormick, N.; Lavalle, C.; Kasanko, M.; Demicheli, L.; Turchini, M.;
Remote Sensing and Data Fusion over Urban Areas, IEEE/ISPRS Joint Works
8-9 Nov. 2001 Page(s):315 - 319
Digital Object Identifier 10.1109/DFUA.2001.985904
AbstractPlus | Full Text: PDF(497 KB) **IEEE CNF**
Rights and Permissions

14. **Interacting with geographic databases: a focus+context approach**
Cicerone, S.; Frigioni, D.; Tarantino, L.;
Database and Expert Systems Applications, 2000. Proceedings. 11th Internati
4-8 Sept. 2000 Page(s):869 - 875
Digital Object Identifier 10.1109/DEXA.2000.875128
AbstractPlus | Full Text: PDF(628 KB) **IEEE CNF**
Rights and Permissions

L10: Entry 33 of 34 File: TDBD Mar 1, 1997

TDB-ACC-NO: NN9703107

DISCLOSURE TITLE: Heirarchical Data Model for a Relational Database based Geographic Information System

PUBLICATION-DATA:
IBM Technical Disclosure Bulletin, March 1997, US

VOLUME NUMBER: 40
ISSUE NUMBER: 3
PAGE NUMBER: 107 - 116

PUBLICATION-DATE: March 1, 1997 (19970301)

CROSS REFERENCE: 0018-8689-40-3-107

DISCLOSURE TEXT:

This document contains drawings, formulas, and/or symbols that will not appear on line. Request hardcopy from ITIRC for complete article. Disclosed is a Relational Data Model for a Geographic Information System (GIS) that supports inheritance. The inheritance model is superior to conventional relational models because it recognizes class relationships that exist in GIS. An implementation of the inheritance model is proposed using relational technology. An actual GIS data model is presented. This model provides support for the three major areas in a GIS: spatial data, application development, and administrative tasks.

L10: Entry 33 of 34 File: TDBD Mar 1, 1997

TDB-ACC-NO: NN9703107

DISCLOSURE TITLE: Heirarchical Data Model for a Relational Database based Geographic Information System

PUBLICATION-DATA:
IBM Technical Disclosure Bulletin, March 1997, US

VOLUME NUMBER: 40
ISSUE NUMBER: 3
PAGE NUMBER: 107 - 116

PUBLICATION-DATE: March 1, 1997 (19970301)

CROSS REFERENCE: 0018-8689-40-3-107

DISCLOSURE TEXT:

This document contains drawings, formulas, and/or symbols that will not appear on line. Request hardcopy from ITIRC for complete article. Disclosed is a Relational Data Model for a Geographic Information System (GIS) that supports inheritance. The inheritance model is superior to conventional relational models because it recognizes class relationships that exist in GIS. An implementation of the inheritance model is proposed using relational technology. An actual GIS data model is presented. This model provides support for the three major areas in a GIS: spatial data, application development, and administrative tasks.

The evolution of GIS, Computer Aided Design (CAD), and Computer Aided Manufacturing (CAM) has required that information be stored and retrieved from non-proprietary databases within an enterprise. This information is very complex. In particular, the commonality between basic tables that model real world objects in GIS/CAD/CAM is difficult to represent in most relational databases. For a GIS, examples of objects include: o A road o A political boundary o A user interface o A user-defined table o System administrative data One of the main obstacles to modelling these tables with relational technology is the ability to recognize that elements of each table are common across many tables. The trend for modelling this commonality has been the adoption of object-oriented databases and software by many GIS/CAD/CAM systems. Businesses have not benefitted from this direction, however, as most enterprises use relational databases. The problem with relational databases is that their table/view mechanism does not provide the flexibility to define a single view across several tables without a relational join. Consider the following: o A road table has "ID", "TYPE", "LOCATION", and "LENGTH" attributes. o A political boundary table has "ID", "TYPE", "LOCATION", "PERIMETER" and "AREA" attributes. What is proposed is: o An Inheritance Data Model for a GIS.

o A mechanism for storage/retrieval of said data model using a relational database. The following represents the basic description of the inheritance data model used in GIS: General Concepts GIS data includes all the enterprise assets (such as parcels, cables, valves, or highway sections) that are modelled and maintained as GIS data, as well as application and operational data (e.g., Macros, Display Rules, User Profiles, and Device Characteristics) used by the GIS software and GIS

applications.

For example, an enterprise may choose to model and maintain land parcel information as GIS data in a GIS. The GIS users and application developers can operate on parcels and the attributes of parcels much as they would access other data in the relational database. In addition, non-GIS users and Corporate applications can access the parcel data through the relational views and application programming interface provided by a GIS.

In a GIS, to determine the appraised value of the parcel owned by John Smith, an user could issue the following query: Select Appraised_Value from Parcel where Owner = 'John Smith' In this example, 'Appraised_Value' and 'Owner' were defined by the application developer or database designer as attributes of Parcel data in the GIS. These attributes are called user defined attributes in the GIS data model. The GIS will also maintain one or more system attributes for each type of GIS data. System attributes are accessed in the same fashion as user-defined attributes in the GIS, for instance, to determine the value (maintained by the GIS) for the area of the parcel owned by John Smith, the following query could be used: Select Area from Parcel where Owner = 'John Smith' Application and Data Independence Under the proposed GIS, the Parcel data discussed in these examples would be physically stored in a relational database. The user-defined attributes, system attributes, and relationships comprising a single Parcel will be distributed among more than one physical table in the relational database. However, the query in the example above is independent of how the Parcel data is physically stored; the user simply operates on Parcel. The concept of application independence from physical storage is important, since it protects applications from the changes to the physical storage format (physical tables). As these changes are made, to provide performance enhancements or to take advantage of enhancements to the underlying relational database, customer applications can be protected. The concept of Views in the relational database is an example of a means of providing application and data independence. A relational view presents a single 'array' of columns that can be operated on in the same fashion as a single physical table, but the columns of the view can represent data that reside in several physical tables, and a column of the view can contain calculated data that is not formally stored in any physical table. Relational applications operating on that view can be protected from changes to the physical tables, in that if the physical table data is re-arranged, the view can be re-created to accommodate the change without having to re-write customer applications that access the view.

GIS Data Model Concepts Relational Views and GIS Object Views In the GIS, an user or application developer can operate on Parcel data much as he would a relational view, although the underlying data for the Parcel is not physically accessed through a single relational table or view. Different components of the parcel data are mapped to a set of columns, as in the case of user and system defined attributes; other components are more complex, such as: picture path elements, edge associations and boundary associations. It is important to note that although the underlying data model for the user's Parcel data can not be mapped to a single set of columns, the GIS permits user access to the Parcel data in a relational fashion as shown in the previous examples. This is done through GIS Object Views. In the GIS data model, an Object View is analogous to a relational view. Users and application developers operate on GIS Object Views much as they would relational views. Object Views provide the data independence for GIS applications that views provide for relational tables. In the previous examples, Parcel is an example of a user-defined Object View created for the storage and maintenance of the enterprise's land parcel data. Several object views can be provided by the GIS, such as Feature, Superspan, Display_Rule, Control_Point, and others. In addition, GIS application and database designers can define their own object views for modelling different kinds of features, tables, and relationships to be maintained by GIS. As mentioned previously, Parcel is an example of a user-defined object view, added to the GIS object views via Nucleus Definition statements.

The various GIS and user-defined object views are related, in that each object view is related to a 'higher level' or more general object view (with the exception of the GIS object view, which is the 'highest' level view). For instance, the user-defined Parcel object view is a classification of the GIS Area Feature object view, which is a classification of the GIS Feature object view. The relationships between all GIS and user-defined object views are explained in "Overview of GIS Object Views". The list of Object View names in the GIS defines the list of names that can be the subject of a From clause in a GIS query (e.g., Select Owner From Parcel). This is similar to identifying all the view names available for manipulating data in a relational database. Once an user or application developer knows the names of these views, queries can be created to select data from the views. Object Views also share the same types of restrictions as relational views, in that relational views can be defined that restrict a user's ability to insert or update data. Likewise, there are GIS object views that are not valid for these operations. All object views are valid for data selection or retrieval (subject to GIS and corporate data security enforcement). Object View Components and Attributes All Object Views in GIS have components and attributes. Components of an Object View can be thought of as complicated attributes (they do not map to a single column of data), whereas an Attribute can be thought of as a 'simple' component (it maps to a single column of data). An object view component can be further broken down into a set, or sets, of attributes.

Projections Once the list of object view names has been identified, it is possible to use the '*' notation to select information from any object view. Once the list of object view components and attributes have been identified, it is possible to select specific sets of values from one or more object views. For instance, once it has been identified that the object view 'Parcel' has two attributes 'Owner' and 'Appraised_Value', users can construct the following queries: Select * From Parcel Where Owner = 'John Smith' Select Owner, Appraised_Value From Parcel Where Owner = 'John Smith' In the GIS, the specification of which component and attribute values are to be returned from a Select operation is called the Projection. An asterisk (*) is used as the notation for the attribute projection of all components and attributes of an object view, just as in SQL, 'Select * from Viewname' will return arrays of values for all columns of the view Viewname. Attribute Evaluation In GIS, any attribute value associated with an instance of data can be evaluated in a where clause. OVERVIEW OF GIS OBJECT VIEWS This section illustrates the relationships between all GIS object views. GIS Object View All data in a GIS database can be accessed via the GIS object view. The components and attributes of the GIS object view are those common to all GIS data. The GIS object view is further classified by the following object views: o Nucleus Object View o Library Object View o Administrative Object View There are two important concepts regarding the set of object views that 'further classify' a particular object view: o Any instance of data that can be accessed via a particular object view can only be accessed via one of the object views that further classify that view. In other words, the list of object views that classify a particular object view are 'mutually exclusive' with regard to instances of data. o The components and attributes of an object view can be treated as components and attributes of any object view that further classifies that view. The GIS object view is illustrated in Fig. 1. Fig. 1 GIS Object View. The GIS Object View is further classified by the three object views as shown in the figure: Administrative, Nucleus, and Library. Put simply, this means that any instance of data that can be accessed using the GIS object view is either nucleus, administrative, or library data. Additional GIS object views are provided to further classify data in the nucleus, administrative, and library object views as indicated by the dotted arrows. Administrative Object View Administrative data is a subset of GIS data that contains information which deals with the creation and management of the collection of DBMS tables that make up the GIS database. This includes definitions for the GIS system and user-defined object views. The "Administrative" object view is further classified by the object views listed in Fig. 2. Administrative object views exclude all library and nucleus data. If an

application developer wanted to select all administrative data, the "Administrative" object view would be used. If the application developer wanted to select only administrative user profile data, the "User" object view could be used. Fig. 2 Administrative Object View. The object views that belong to the set of administrative object views are illustrated above. All instances of administrative data belong to one of these object views. There are no further object views or further classifications of object views for administrative data. Library Object View Library Objects consist of various rules and parameters that control how the user interacts with the system, how information is captured, how information is graphically and textually realized, and the meaning of units and projections. In short, this information defines much of the operation of a GIS and may be specific to individual users of groups of users or may be generic across an entire enterprise. The "Library" object view is further classified by the object views listed in Fig. 3. Library object views exclude all nucleus and administrative data. If an application developer wanted to select all library data, the "Library" object view would be used. If an application developer wanted to select only Macros, the "Macro" object view could be used.

Fig. 3 Library Object View. The object views that belong to the set of library object views are illustrated above. All instances of library data belong to one of these object views. There are no further object views or further classifications of object views for library data. Nucleus Object View Nucleus data is a subset of GIS data that contains geographic data and attributes associated with that data and is the subset of GIS data used to model and maintain an enterprise's geographic assets.

The following characteristics distinguish nucleus data from administrative or library data in GIS: o Nucleus data can be evaluated based on spatial (or location) characteristics. o Nucleus data can be manipulated by check-in, check-out, load-for-update and extract-updates operations. o The Nucleus object views can be extended by the database designer or application developer through GIS Nucleus definitions. Nucleus definitions permit the creation of user-defined object views for modelling features, relationships, and user-tables. Nucleus object views exclude all library and administrative data. If an application developer wanted to select all nucleus data, then the "Nucleus" object view would be used. If an application developer wanted to select only Features, then the "Feature" object view would be used. The Feature, Relationship and User_Table object views are further classified by additional GIS and user-defined object views. These object views are described in the following sections: Feature Object View The following characteristics distinguish feature data from other data in GIS: o Features can be associated with Display Rules.

Features have the following location characteristics: o Absolute Point o Node o Extent o Feature_Path The "Feature" object views exclude all library data and administrative data and all nucleus data such as user tables, boundaries, polygon sets, etc.. If an application developer wanted to select all feature data, then the "Feature" object view would be used. If the application developer wanted to select only point features, the "Point" object view would be used. If the application developer wanted to select only those point features defined as poles having a feature class of "telephone_pole", then the "telephone_pole" object view would be used. Each of the Feature object views may be further classified by user-defined feature names, such as 'Parcel' or 'Highway', that are created via GIS Nucleus Definitions. Nucleus definitions provide a means of associating user-defined attributes, such as 'Owner' with the Parcel object view. Each object view created via Feature definitions may be further classified, if desired, by specifying one or more Feature Classes within the Feature Definition. For instance, the enterprise may choose to further classify 'Parcel' features by defining 'Rural_Lot', 'Municipal_Lot' feature classes. Any user-defined attributes associated with Parcel data are also associated with all feature classes of the Parcel data; however, different Display Rules may be associated with the various feature classes for

Parcel data (e.g. 'Rural_Lot' features can be rendered differently than 'Municipal_Lot' features).

If an application developer or database designer creates a feature definition for an Area feature named Parcel, having a user-defined attribute named 'Owner', and feature classes 'Rural_Lot', 'Municipal_Lot', the following queries could be used to manipulate Parcel data: Select * from Parcel where Owner = 'John Smith' Select * from Rural_Lot where Owner = 'John Smith' Select * from Municipal_Lot where Owner = 'John Smith' The first query will return all instances of parcels owned by 'John Smith', the second will return only those instances of parcels that are classified as rural lots, the third, only municipal lots.

User_Table Object View User tables provide the ability to model non-geographic data (sets of attributes) as part of the GIS nucleus. The instances of user table data can be related to features via GIS user-defined relationships, and the user table data can be manipulated with checkin, checkout, load and extract operations. User tables are defined via GIS Nucleus definitions. Each User_Table name, when defined by the application developer or database designer, becomes an object view which further classifies the User_Table object view.

Relationship Object View The Relationship object view is further classified by GIS system maintained relationships, such as Layer_assoc and Theme_assoc. Application developers and database designers can create user-defined relationships via GIS Nucleus definitions. GIS provides a means of creating user-defined relationships. Each user-defined relationship name, when defined by the application developer or database designer, becomes an object view which further classifies the Relationship object view. As with features and user tables, the GIS nucleus definitions provide a means of associating sets of (user-defined) attributes with user-defined relationships. Object View Considerations for Applications and Data Access Recall that the list of GIS Object Views defines the 'names' that can be the subject of a From clause to access GIS data. For instance, using the Object Views listed thus far, the following are valid queries: A) Select * From GIS B) Select * From Nucleus C) Select * From Feature D) Select * From Span If a GIS database designer creates a Feature Definition for 'Highway', having a Feature Class called 'Interstate', then the following two queries would also be valid:

E) Select * From Highway F) Select * From Interstate The relationship between these object views is much like a relationship between various views on the same set of relational tables. The object view specified in the From clause defines the scope of the data that can possibly be returned from the query. The object view is also a consideration in performance of queries, since the more general object views will generally indicate searches through a wider range of data. Conclusion This disclosure presented a Relational Data Model for a Geographic Information System that supports inheritance. The inheritance model is superior to conventional relational models, because it recognizes the inheritance relationships that exist in GIS. An implementation of the inheritance model was proposed using relational technology. A relational approach has several advantages over object-oriented implementations. First, relational DBMSs are commonplace in the business world. Second, relational systems support application and data independence through abstraction and views. Third, distributed relational systems are reaching maturity. An actual GIS data model was presented. This model provided support for the three major areas in a GIS: spatial data, application development, and administrative tasks.

Previous Doc. Next Doc Go to Doc#

L10: Entry 33 of 34 File: TDBD Mar 1, 1997

TDB-ACC-NO: NN9703107

DISCLOSURE TITLE: Heirarchical Data Model for a Relational Database based Geographic Information System

PUBLICATION-DATA:
IBM Technical Disclosure Bulletin, March 1997, US

VOLUME NUMBER: 40
ISSUE NUMBER: 3
PAGE NUMBER: 107 - 116

PUBLICATION-DATE: March 1, 1997 (19970301)

CROSS REFERENCE: 0018-8689-40-3-107

DISCLOSURE TEXT:

This document contains drawings, formulas, and/or symbols that will not appear on line. Request hardcopy from ITIRC for complete article. Disclosed is a Relational Data Model for a Geographic Information System (GIS) that supports inheritance. The inheritance model is superior to conventional relational models because it recognizes class relationships that exist in GIS. An implementation of the inheritance model is proposed using relational technology. An actual GIS data model is presented. This model provides support for the three major areas in a GIS: spatial data, application development, and administrative tasks.

First Hit Previous Doc Next Doc Go to Doc#

Generate Collection Print

L10: Entry 33 of 34 File: TDBD Mar 1, 1997

TDB-ACC-NO: NN9703107

DISCLOSURE TITLE: Heirarchical Data Model for a Relational Database based Geographic Information System

PUBLICATION-DATA:
IBM Technical Disclosure Bulletin, March 1997, US

VOLUME NUMBER: 40
ISSUE NUMBER: 3
PAGE NUMBER: 107 - 116

PUBLICATION-DATE: March 1, 1997 (19970301)

CROSS REFERENCE: 0018-8689-40-3-107

DISCLOSURE TEXT:

This document contains drawings, formulas, and/or symbols that will not appear on line. Request hardcopy from ITIRC for complete article. Disclosed is a Relational Data Model for a Geographic Information System (GIS) that supports inheritance. The inheritance model is superior to conventional relational models because it recognizes class relationships that exist in GIS. An implementation of the inheritance model is proposed using relational technology. An actual GIS data model is presented. This model provides support for the three major areas in a GIS: spatial data, application development, and administrative tasks.

The evolution of GIS, Computer Aided Design (CAD), and Computer Aided Manufacturing (CAM) has required that information be stored and retrieved from non-proprietary databases within an enterprise. This information is very complex. In particular, the commonality between basic tables that model real world objects in GIS/CAD/CAM is difficult to represent in most relational databases. For a GIS, examples of objects include: o A road o A political boundary o A user interface o A user-defined table o System administrative data One of the main obstacles to modelling these tables with relational technology is the ability to recognize that elements of each table are common across many tables. The trend for modelling this commonality has been the adoption of object-oriented databases and software by many GIS/CAD/CAM systems. Businesses have not benefitted from this direction, however, as most enterprises use relational databases. The problem with relational databases is that their table/view mechanism does not provide the flexibility to define a single view across several tables without a relational join. Consider the following: o A road table has "ID", "TYPE", "LOCATION", and "LENGTH" attributes. o A political boundary table has "ID", "TYPE", "LOCATION", "PERIMETER" and "AREA" attributes. What is proposed is: o An Inheritance Data Model for a GIS.

o A mechanism for storage/retrieval of said data model using a relational database. The following represents the basic description of the inheritance data model used in GIS: General Concepts GIS data includes all the enterprise assets (such as parcels, cables, valves, or highway sections) that are modelled and maintained as GIS data, as well as application and operational data (e.g., Macros, Display Rules, User Profiles, and Device Characteristics) used by the GIS software and GIS

applications.

For example, an enterprise may choose to model and maintain land parcel information as GIS data in a GIS. The GIS users and application developers can operate on parcels and the attributes of parcels much as they would access other data in the relational database. In addition, non-GIS users and Corporate applications can access the parcel data through the relational views and application programming interface provided by a GIS.

In a GIS, to determine the appraised value of the parcel owned by John Smith, an user could issue the following query: Select Appraised_Value from Parcel where Owner = 'John Smith' In this example, 'Appraised_Value' and 'Owner' were defined by the application developer or database designer as attributes of Parcel data in the GIS. These attributes are called user defined attributes in the GIS data model. The GIS will also maintain one or more system attributes for each type of GIS data. System attributes are accessed in the same fashion as user-defined attributes in the GIS, for instance, to determine the value (maintained by the GIS) for the area of the parcel owned by John Smith, the following query could be used: Select Area from Parcel where Owner = 'John Smith' Application and Data Independence Under the proposed GIS, the Parcel data discussed in these examples would be physically stored in a relational database. The user-defined attributes, system attributes, and relationships comprising a single Parcel will be distributed among more than one physical table in the relational database. However, the query in the example above is independent of how the Parcel data is physically stored; the user simply operates on Parcel. The concept of application independence from physical storage is important, since it protects applications from the changes to the physical storage format (physical tables). As these changes are made, to provide performance enhancements or to take advantage of enhancements to the underlying relational database, customer applications can be protected. The concept of Views in the relational database is an example of a means of providing application and data independence. A relational view presents a single 'array' of columns that can be operated on in the same fashion as a single physical table, but the columns of the view can represent data that reside in several physical tables, and a column of the view can contain calculated data that is not formally stored in any physical table. Relational applications operating on that view can be protected from changes to the physical tables, in that if the physical table data is re-arranged, the view can be re-created to accommodate the change without having to re-write customer applications that access the view.

GIS Data Model Concepts Relational Views and GIS Object Views In the GIS, an user or application developer can operate on Parcel data much as he would a relational view, although the underlying data for the Parcel is not physically accessed through a single relational table or view. Different components of the parcel data are mapped to a set of columns, as in the case of user and system defined attributes; other components are more complex, such as: picture path elements, edge associations and boundary associations. It is important to note that although the underlying data model for the user's Parcel data can not be mapped to a single set of columns, the GIS permits user access to the Parcel data in a relational fashion as shown in the previous examples. This is done through GIS Object Views. In the GIS data model, an Object View is analogous to a relational view. Users and application developers operate on GIS Object Views much as they would relational views. Object Views provide the data independence for GIS applications that views provide for relational tables. In the previous examples, Parcel is an example of a user-defined Object View created for the storage and maintenance of the enterprise's land parcel data. Several object views can be provided by the GIS, such as Feature, Superspan, Display_Rule, Control_Point, and others. In addition, GIS application and database designers can define their own object views for modelling different kinds of features, tables, and relationships to be maintained by GIS. As mentioned previously, Parcel is an example of a user-defined object view, added to the GIS object views via Nucleus Definition statements.

The various GIS and user-defined object views are related, in that each object view is related to a 'higher level' or more general object view (with the exception of the GIS object view, which is the 'highest' level view). For instance, the user-defined Parcel object view is a classification of the GIS Area Feature object view, which is a classification of the GIS Feature object view. The relationships between all GIS and user-defined object views are explained in "Overview of GIS Object Views". The list of Object View names in the GIS defines the list of names that can be the subject of a From clause in a GIS query (e.g., Select Owner From Parcel). This is similar to identifying all the view names available for manipulating data in a relational database. Once an user or application developer knows the names of these views, queries can be created to select data from the views. Object Views also share the same types of restrictions as relational views, in that relational views can be defined that restrict a user's ability to insert or update data. Likewise, there are GIS object views that are not valid for these operations. All object views are valid for data selection or retrieval (subject to GIS and corporate data security enforcement). Object View Components and Attributes All Object Views in GIS have components and attributes. Components of an Object View can be thought of as complicated attributes (they do not map to a single column of data), whereas an Attribute can be thought of as a 'simple' component (it maps to a single column of data). An object view component can be further broken down into a set, or sets, of attributes.

Projections Once the list of object view names has been identified, it is possible to use the '*' notation to select information from any object view. Once the list of object view components and attributes have been identified, it is possible to select specific sets of values from one or more object views. For instance, once it has been identified that the object view 'Parcel' has two attributes 'Owner' and 'Appraised_Value', users can construct the following queries: Select * From Parcel Where Owner = 'John Smith' Select Owner, Appraised_Value From Parcel Where Owner = 'John Smith' In the GIS, the specification of which component and attribute values are to be returned from a Select operation is called the Projection. An asterisk (*) is used as the notation for the attribute projection of all components and attributes of an object view, just as in SQL, 'Select * from Viewname' will return arrays of values for all columns of the view Viewname. Attribute Evaluation In GIS, any attribute value associated with an instance of data can be evaluated in a where clause. OVERVIEW OF GIS OBJECT VIEWS This section illustrates the relationships between all GIS object views. GIS Object View All data in a GIS database can be accessed via the GIS object view. The components and attributes of the GIS object view are those common to all GIS data. The GIS object view is further classified by the following object views: o Nucleus Object View o Library Object View o Administrative Object View There are two important concepts regarding the set of object views that 'further classify' a particular object view: o Any instance of data that can be accessed via a particular object view can only be accessed via one of the object views that further classify that view. In other words, the list of object views that classify a particular object view are 'mutually exclusive' with regard to instances of data. o The components and attributes of an object view can be treated as components and attributes of any object view that further classifies that view. The GIS object view is illustrated in Fig. 1. Fig. 1 GIS Object View. The GIS Object View is further classified by the three object views as shown in the figure: Administrative, Nucleus, and Library. Put simply, this means that any instance of data that can be accessed using the GIS object view is either nucleus, administrative, or library data. Additional GIS object views are provided to further classify data in the nucleus, administrative, and library object views as indicated by the dotted arrows. Administrative Object View Administrative data is a subset of GIS data that contains information which deals with the creation and management of the collection of DBMS tables that make up the GIS database. This includes definitions for the GIS system and user-defined object views. The "Administrative" object view is further classified by the object views listed in Fig. 2. Administrative object views exclude all library and nucleus data. If an

application developer wanted to select all administrative data, the "Administrative" object view would be used. If the application developer wanted to select only administrative user profile data, the "User" object view could be used. Fig. 2 Administrative Object View. The object views that belong to the set of administrative object views are illustrated above. All instances of administrative data belong to one of these object views. There are no further object views or further classifications of object views for administrative data. Library Object View Library Objects consist of various rules and parameters that control how the user interacts with the system, how information is captured, how information is graphically and textually realized, and the meaning of units and projections. In short, this information defines much of the operation of a GIS and may be specific to individual users of groups of users or may be generic across an entire enterprise. The "Library" object view is further classified by the object views listed in Fig. 3. Library object views exclude all nucleus and administrative data. If an application developer wanted to select all library data, the "Library" object view would be used. If an application developer wanted to select only Macros, the "Macro" object view could be used.

Fig. 3 Library Object View. The object views that belong to the set of library object views are illustrated above. All instances of library data belong to one of these object views. There are no further object views or further classifications of object views for library data. Nucleus Object View Nucleus data is a subset of GIS data that contains geographic data and attributes associated with that data and is the subset of GIS data used to model and maintain an enterprise's geographic assets.

The following characteristics distinguish nucleus data from administrative or library data in GIS: o Nucleus data can be evaluated based on spatial (or location) characteristics. o Nucleus data can be manipulated by check-in, check-out, load-for-update and extract-updates operations. o The Nucleus object views can be extended by the database designer or application developer through GIS Nucleus definitions. Nucleus definitions permit the creation of user-defined object views for modelling features, relationships, and user-tables. Nucleus object views exclude all library and administrative data. If an application developer wanted to select all nucleus data, then the "Nucleus" object view would be used. If an application developer wanted to select only Features, then the "Feature" object view would be used. The Feature, Relationship and User_Table object views are further classified by additional GIS and user-defined object views. These object views are described in the following sections: Feature Object View The following characteristics distinguish feature data from other data in GIS: o Features can be associated with Display Rules.

Features have the following location characteristics: o Absolute Point o Node o Extent o Feature_Path The "Feature" object views exclude all library data and administrative data and all nucleus data such as user tables, boundaries, polygon sets, etc.. If an application developer wanted to select all feature data, then the "Feature" object view would be used. If the application developer wanted to select only point features, the "Point" object view would be used. If the application developer wanted to select only those point features defined as poles having a feature class of "telephone_pole", then the "telephone_pole" object view would be used. Each of the Feature object views may be further classified by user-defined feature names, such as 'Parcel' or 'Highway', that are created via GIS Nucleus Definitions. Nucleus definitions provide a means of associating user-defined attributes, such as 'Owner' with the Parcel object view. Each object view created via Feature definitions may be further classified, if desired, by specifying one or more Feature Classes within the Feature Definition. For instance, the enterprise may choose to further classify 'Parcel' features by defining 'Rural_Lot', 'Municipal_Lot' feature classes. Any user-defined attributes associated with Parcel data are also associated with all feature classes of the Parcel data; however, different Display Rules may be associated with the various feature classes for

Parcel data (e.g. 'Rural_Lot' features can be rendered differently than 'Municipal_Lot' features).

If an application developer or database designer creates a feature definition for an Area feature named Parcel, having a user-defined attribute named 'Owner', and feature classes 'Rural_Lot', 'Municipal_Lot', the following queries could be used to manipulate Parcel data: Select * from Parcel where Owner = 'John Smith' Select * from Rural_Lot where Owner = 'John Smith' Select * from Municipal_Lot where Owner = 'John Smith' The first query will return all instances of parcels owned by 'John Smith', the second will return only those instances of parcels that are classified as rural lots, the third, only municipal lots.

User_Table Object View User tables provide the ability to model non-geographic data (sets of attributes) as part of the GIS nucleus. The instances of user table data can be related to features via GIS user-defined relationships, and the user table data can be manipulated with checkin, checkout, load and extract operations. User tables are defined via GIS Nucleus definitions. Each User_Table name, when defined by the application developer or database designer, becomes an object view which further classifies the User_Table object view.

Relationship Object View The Relationship object view is further classified by GIS system maintained relationships, such as Layer_assoc and Theme_assoc. Application developers and database designers can create user-defined relationships via GIS Nucleus definitions. GIS provides a means of creating user-defined relationships. Each user-defined relationship name, when defined by the application developer or database designer, becomes an object view which further classifies the Relationship object view. As with features and user tables, the GIS nucleus definitions provide a means of associating sets of (user-defined) attributes with user-defined relationships. Object View Considerations for Applications and Data Access Recall that the list of GIS Object Views defines the 'names' that can be the subject of a From clause to access GIS data. For instance, using the Object Views listed thus far, the following are valid queries: A) Select * From GIS B) Select * From Nucleus C) Select * From Feature D) Select * From Span If a GIS database designer creates a Feature Definition for 'Highway', having a Feature Class called 'Interstate', then the following two queries would also be valid:

E) Select * From Highway F) Select * From Interstate The relationship between these object views is much like a relationship between various views on the same set of relational tables. The object view specified in the From clause defines the scope of the data that can possibly be returned from the query. The object view is also a consideration in performance of queries, since the more general object views will generally indicate searches through a wider range of data. Conclusion This disclosure presented a Relational Data Model for a Geographic Information System that supports inheritance. The inheritance model is superior to conventional relational models, because it recognizes the inheritance relationships that exist in GIS. An implementation of the inheritance model was proposed using relational technology. A relational approach has several advantages over object-oriented implementations. First, relational DBMSs are commonplace in the business world. Second, relational systems support application and data independence through abstraction and views. Third, distributed relational systems are reaching maturity. An actual GIS data model was presented. This model provided support for the three major areas in a GIS: spatial data, application development, and administrative tasks.

Previous Doc Next Doc Go to Doc#

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 837 341 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
22.04.1998 Bulletin 1998/17

(51) Int. Cl.[6]: G01S 5/14, H04Q 7/38, G08G 1/127

(21) Application number: 97203380.7

(22) Date of filing: 01.09.1995

(84) Designated Contracting States:
BE CH DE DK ES FR GB IT LI NL SE

(30) Priority: 01.09.1994 GB 9417600

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
95930610.1 / 0 777 863

(71) Applicant:
BRITISH TELECOMMUNICATIONS public limited company
London EC1A 7AJ (GB)

(72) Inventors:
- Mannings, Robin Thomas
  Ipswich, Suffolk, IP5 7TU (GB)
- Wall, Nigel David Charles
  Ipswich, Suffolk, IP4 2TL (GB)

(74) Representative:
Lidbetter, Timothy Guy Edwin
BT Group Legal Services,
Intellectual Property Department,
8th Floor, Holborn Centre,
120 Holborn
London EC1N 2TE (GB)

Remarks:
This application was filed on 31 - 10 - 1997 as a divisional application to the application mentioned under INID code 62.

(54) Navigation information system

(57) A navigation information system comprises a communications system having a fixed part (11 to 20) and at least one mobile part (1 to 10), the fixed part including a data storage and processing means 15 for identifying the location of a mobile unit, generating guidance information appropriate to that location and transmitting it to the mobile unit. The data exchange is by means of DTMF codes. By locating most of the complexity with the service provider, in particular the navigation computer 15 and geographical database 17, the system can be readily updated and the capital cost of the in-vehicle system, which in its simplest form may be a standard cellular telephone 1, can be minimised. The user makes a request for guidance information, and the system, having determined the user's present location, then transmits instructions to the user. The user's present location can be determined by means such as a Satellite Positioning System 7.

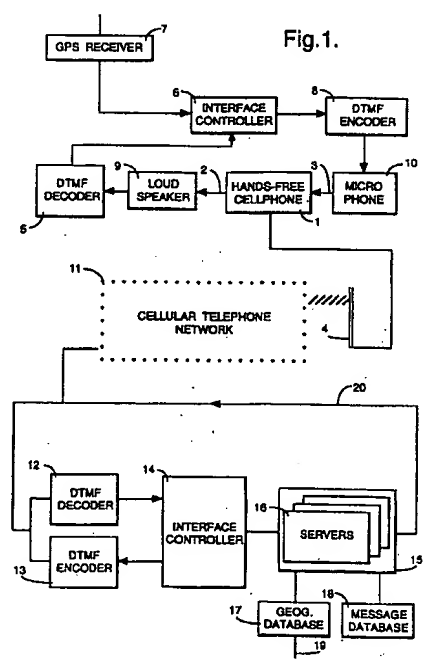


Fig.1.

EP 0 837 341 A1

Description

This invention relates to navigation information systems. It is particularly suitable for use in providing users of road vehicles with route guidance, but other applications are possible and are discussed below.

Navigation of a vehicle through an unfamiliar complex road network is a difficult task. Large amounts of fuel and time are wasted as a result of drivers getting lost or using an inefficient route. Accidents can also be caused by drivers attempting to read maps or complex road signs and losing concentration on the road ahead. Moreover, a driver may choose an inefficient route as a result of using an out-of-date map.

An additional problem can occur even if a driver knows a route to his or her destination. That route may be congested or blocked as a result of accidents or maintenance work, so that an alternative route would be more efficient.

Several proposals have been made for navigation guidance systems. In some such proposals a vehicle-borne system has a navigation computer and a geographical information system which is essentially a digitised map stored on a CD-ROM. The system gives the driver information and guidance by screen and/or speech display. These systems would be very expensive. Each vehicle requires a navigation computer and geographical information system. The cost of the complex vehicle-borne equipment involved is estimated to be in the region of £1000. The system is complex to operate, and could only be safely operated by the driver whilst the vehicle is stationary. The geographical information system would require periodic updating, which requires new disks to be distributed to subscribers from time to time.

In some proposed systems of this type real-time data would be broadcast over a radio network to update fixed information held on the geographical information system. Even so, the geographical information system would only be accurate up to its last update. Moreover, a broadcast channel needs to be allocated for the updating service.

It has also been proposed that the guidance service provider collects statistical traffic flow data from which traffic congestion predictions can be made which are fed into the real-time data to be broadcast. The traffic flow data may be collected using roadside sensors, or they may be collected by monitoring the operation of the mobile user equipment. The latter approach can only collect data relating to users of the system, but has a lower capital cost.

In an alternative approach a system of short-range roadside beacons is used to transmit guidance information to passing vehicles equipped with simple transceivers. The beacons transmit information to suitably equipped passing vehicles to give turn instructions appropriate to their chosen destinations. For each beacon the territory to be covered is divided into as many zones as there are exits from the junction the beacon relates to. The zone in which the user's chosen destination falls is determined, and instructions are given appropriate to that zone. At any given beacon all vehicles whose destinations are in the same zone get the same instruction. The definitions of the zones are dependant on the location of the beacons, and each zone comprises the set of destinations which should be reached from the beacon by taking the direction associated with that zone.

Each beacon only gives instructions for reaching the next beacon along the route to the vehicle's destination. For two vehicles starting from the same point for different destinations for which the routes are initially coincident, the beacons along the coincident section of route will each give both users the same instructions, because for those beacons both users are travelling to the same zone. Only for the beacon at the point of divergence are the two users' destinations in different zones, and therefore different instructions are given.

The beacons' programming may be modified from time to time by control signals from a central control station, in a way analogous to remotely controlled adjustable signposts, but in its interactions with the user equipment the beacon is autonomous, identifying which of its zones the user's desired destination is in, and transmitting the appropriate "turn" information to get it to the next beacon on the way. The beacon has no knowledge of the rest of the route.

Each beacon has a detailed map of a small local area (the boundaries of which are, in fact, the adjacent beacons), and if the destination is in this area the beacon gives full information of the route to the destination. The system can therefore provide a user with directions to a destination defined more precisely than the beacon spacing. However, at the beginning of a journey, a user cannot use the system until he encounters a beacon.

This proposed system allows instant updating of the guidance instructions from a central control, and simpler in-vehicle equipment, but requires vast capital expenditure in roadside beacons.

A problem encountered with both the proposed systems described above is that it is difficult for them to provide alternative routings in response to congestion, either current or future, without the risk of creating worse problems on the alternative routes. Although predictions of regularly occurring congestion peaks are relatively simple to programme into the guidance information, and, at least in the beacon system, real-time updates on road congestion can also be fed to the programming of the beacons, the control system does not have any information of vehicle movements from which to predict future congestion. In any case, if the system is in use by a significant fraction of the vehicles, the system will tend to produce congestion on the diversionary routes.

According to a first aspect of the invention, there is provided a navigation information system for providing

information to one or more mobile users dependent on their locations, the system comprising:

means for determining the location of a mobile unit requesting guidance information,
means for generating information for guidance of the user of the mobile unit according to the present location of the mobile unit,
and a communications system for transmitting the guidance information so generated to the mobile unit,

wherein the navigation information system has means for communicating with the mobile unit using dual tone multi-frequency signals.

According to a second aspect, there is provided a mobile part of a navigation information system, comprising means for identifying the present position of the mobile unit, means for transmitting to a co-opertaing fixed part, over a communications link, information relating to the present location of the mobile unit, and guidance instruction means controllable by guidance instruction information received over the communications link, whereby guidance instructions related to the present location can be communicated to a user by means of the guidance instruction means,
characterised in that the mobile unit has means for communicating with the co-operating fixed part using dual tone multi-frequency (DTMF) signals.

According to a third aspect, there is provided a method of providing navigation guidance information to mobile units of a mobile radio system, the information being dependent on the locations of the mobile units, the method comprising the steps of:

- transmitting, from a mobile unit to a fixed part, a request for navigation guidance;
- determining the location of the mobile unit;
- generating guidance information on the basis of the location information and navigation data stored in the fixed part; and
- transmitting the guidance information from the fixed part to the mobile unit;

characterised in that at least part of the data transmitted to and/or from the mobile unit is in the form of DTMF codes.

For an analogue cellular radio network DTMF is an ideal signalling medium when only short status messages are required to be transmitted. It can survive in the severe signal fading and noise of the mobile environment which frequently precludes the use of fast phase or frequency shift data modulation. Another advantage is the ability to co-exist with speech. For example a DTMF data burst containing vehicle position data could be sent at the start of a call and at intervals during the call. Other simple coded DTMF messages can also be conveyed to indicate emergencies, provide simple driver indications (e.g. illuminated arrows to turn left or right) or trigger synthetic speech generated by another sub-system in the vehicle.

This invention has advantages over both the prior art systems discussed above. Considerable improvements can be made over the prior on-board navigation system proposals by putting the intelligence in the fixed part of the system. Firstly, there is no need to distribute maps or updates to subscribers because the data is held centrally. New roads can be added to the system at the instant they are opened. Total capital expenditure is minimised since all users share the same database. Moreover, the computing resources are used more efficiently, because an in-vehicle system spends most of its time inactive but a centralised system can be time-shared.

Moreover, in contrast to the prior art roadside beacon system, the invention can be implemented with little deployment of equipment in the field, thereby offering considerable economies in capital cost and maintenance, and allowing rapid installation and modification of the system to meet changing requirements.

Preferably the system includes means for determining the location of the mobile part in relation to a geographical overlay comprising a plurality of overlay areas, and means for transmitting information associated with an overlay area which includes the location of the mobile part, whereby a mobile part within that overlay area receives information associated with that overlay area. This allows information associated with a particular overlay area to be transmitted to any mobile units in that overlay area. The system may also comprise means for determining when a mobile part enters a predetermined overlay area, and means for transmitting a message, to a user other than the said mobile part, in response to the said mobile part entering the predetermined overlay area. For example, one overlay area may cover part of a road approaching a junction, and the message may be the appropriate instruction to the driver, as he approaches the junction, as to which way he should turn. Each individual overlay area therefore gives navigation instructions specific to that overlay area. The overlay areas may overlap, and may be of any size down to the practical minimum of the resolution of the location determination process. Large overlay areas are suitable for transmitting general information, whilst smaller areas can be used to target information to users in very precise locations, such as individual elements of a complicated road layout. The overlay areas may be delimited in two or three dimensions.

An advantage of this preferred arrangement over the fixed beacon systems is that the geographical overlay can be readily modified. Advantageously, the system includes means for storing a digital representation of the geographical overlay, and means for modifying the stored representation such that the configurations of the overlay areas may be selected to meet changing requirements. The overlay areas can be readily com-

bined or subdivided, or their boundaries otherwise altered to meet changing circumstances without any modification to the hardware, simply by reconfiguring the geographical overlay defined in the central database. Moreover, unlike the prior art beacon system discussed above, there is no major cost in street furniture and supporting infrastructure, because existing cellular mobile communications systems may be used to transmit the instructions from a central database. If the driver enters an overlay area which is not on the route chosen by the system, an error message can be transmitted. Such messages may be transmitted to a user other than the mobile unit, for instance in order to monitor the whereabouts of valuable cargoes or of personnel working away from a base.

The geographical overlay may also be used to operate an access-control system, for example for site security or for levying tolls. In this arrangement, if a user enters an overlay area for which he does not have permission, an alert signal can be sent to a system controller, or to security staff on site who can intercept the interloper. Means may be provided (either in a fixed location or with the mobile user) to store a value associated with the mobile unit, and means arranged to modify the stored value in response to the messages transmitted in accordance with the location of the mobile unit, either to increment the value e.g. for subsequent billing, or to decrement the value e.g. in a prepaid stored-value device.

The fixed part may include means for storing map information or other data for use in providing information, herein referred to as guidance data, means for updating the stored guidance data, means for identifying mobile parts to which the updated data are applicable, and means for transmitting such data over the communications system to the mobile parts so identified. This allows information about changing traffic situations to be transmitted to all users who will be affected, without needing to broadcast the details to other users as would be the case with those prior art systems where updating is possible.

Although the information transmitted to the user is specific to the location, information about the user can be processed centrally. This allows short-term traffic predictions to be made. The guidance data transmitted to the mobile units can therefore be based on the position measurements of a plurality of the mobile parts. If the mobile parts are vehicles, these position measurements will identify the locations of roads, and an indication of their traffic density. As new roads are built or routes are diverted, traffic will move to the new routes. Measuring the position of the traffic will therefore result in the data being updated automatically. To reduce the volume of information transmitted, the fixed part may comprise means for transmitting to the mobile part an expected range of movement information and for receiving from the mobile part movement measurements outside the expected range, and the mobile part comprising means for measuring location and time to derive movement information, means to compare the movement information with the expected range received from a fixed part of the system, and means to automatically report to the fixed system movement measurements outside the expected range. In this way only exceptional traffic conditions are reported.

The fixed part may include means for generating and maintaining guidance data based on vehicle movement data derived from time information and position measurements of a plurality of the mobile parts and/or estimations of future locations of the mobile parts based on the guidance information previously transmitted to the mobile parts. Estimations of future locations of the mobile parts based on the guidance information previously transmitted to the mobile parts can be used to make estimates of future traffic situations.

The data stored in the data storage means may be updated, for example in response to changing traffic conditions, accidents, or highway maintenance. The system may include means for identifying the mobile units to which the updated data are applicable, and transmitting amended instructions over the communications system to said mobile parts. With knowledge of the journeys being planned by a large number of users, a better prediction of demand for particular roads (and hence of congestion on those roads) can be built up. This can be more stable than existing autonomous route-planning systems because the navigation system can take account of the journeys planned for other users.

Advantageously the invention can be implemented using a public cellular radio data service on an individual dial-up basis, providing a simple mechanism for billing and avoiding the need for a separate radio transmission system.

The means for determining the location of the mobile part may comprise means to interrogate a location-identifying means forming part of the mobile part operating for example by means of dead reckoning from a known start point, using an inertial navigation system or distance and direction measuring devices such as a compass and an odometer. Alternatively, the means for locating position may include means for identifying the location of the mobile part in relation to elements of the fixed part of the communications system. The location of the mobile part may be determined by a radio location system associated with the cellular radio system. In another alternative arrangement, a satellite navigation system may be used. In one preferred arrangement the fixed part has means to determine the approximate location of the mobile part, and the location identifying means of the mobile part is arranged to respond to a location request from the interrogation means with a non-unique location signal which, in combination with the approximate location determined by the fixed part, determines a unique location.

In a preferred arrangement, the fixed part and the

mobile parts each have a satellite navigation system receiver, and the positions of the mobile parts as measured by the satellite navigation system are compared with those of the fixed part as measured by the satellite navigation system. The position of the fixed part can be known with great accuracy and provides a reference measurement which allows the position of the mobile part to be determined with greater accuracy than is possible by direct measurement using the satellite system alone.

Preferably the fixed part has one or more servers and means for allocating a server to a mobile part only when it requires service. In practice only a very small number of mobile units will require service at any given time, so this allows the computing resources of the fixed part to be used most efficiently, and the system can support many more mobile units in total than it has server capacity for. This is in contrast to the prior art system discussed above, in which each mobile unit requires a dedicated computer carried on board, which is only used for a fraction of the time. Moreover, all the servers can use a common road-use database, which can use the information on routes it has planned for mobile users to build a prediction of future road use status, such as likely congestion points, and build this into its guidance instruction process. For example the system can be arranged such that it does not direct more than a predetermined number of users to use a particular stretch of road at a particular time, and finds alternative routes for any users who would otherwise be directed along that road at that time. In this way the system can predict likely congestion points and take pre-emptive action.

The mobile part may include guidance instruction means controllable by instructions contained in the guidance information transmitted from the fixed part over the communications link, whereby guidance instructions can be communicated to the user by means of the guidance instruction means.

For some applications the vehicle may be controlled directly in response to the guidance information received over the communications link. However, for use on the public highway, it is preferable that the guidance information controls display means, which may be visual or audible or both, to indicate to a driver the direction to take.

The guidance instruction means may be programmable from the fixed part over the communications link, either automatically or by a human operator. The guidance instruction means may include a speech synthesiser, which may be located in the fixed part, transmitting voice messages to the user over the communications system, or may be located in the mobile unit and controlled by data messages from the fixed part. The former arrangement allows the mobile unit to be simplified, whilst the latter arrangement requires a smaller signalling load.

In the described embodiment the mobile part is in a vehicle, but it may be a hand-held device for guiding a pedestrian. In one form, the mobile part may be a conventional mobile cellular radio unit. This allows a basic service to be provide to a user without the need for any dedicated equipment.

Embodiments of the invention will now be described by way of example with reference to the drawings, in which:

Figure 1 shows a mobile part and a fixed part of a navigation information system according to an embodiment of the invention;

Figure 2 illustrates how the invention may be applied to a simple road layout;

Figure 3 illustrates the division of a territory into zones according to the instructions generated by the system;

Figure 4 illustrates an application of the invention to a more complex road layout;

Figures 5a and 5b illustrate the modification of an overlay in response to a change in traffic circumstances; and

Figure 6 illustrates a road network, showing overlay areas defined by the method of the invention in relation to a cellular radio network

According to the embodiment of Figure 1 the navigation system has a fixed part (comprising elements 12 to 19) and a number of mobile parts, of which one only is shown (comprising elements 1 to 10), interconnected by a cellular telephone network 11.

The mobile part comprises a mobile telephone 1 having an audio output 2, an audio input 3 and a radio antenna (transmit/receive) 4. The output 2 is connected to a decoder 5 to translate Dual-Tone Multi-Frequency (DTMF) signals received by the telephone 1 into data which is fed to an interface controller 6. The interface controller 6 also receives input from a GPS (Global Positioning System) satellite receiver 7. The interface controller transmits data to a DTMF encoder 8 which generates tones to be fed to the audio input of the mobile telephone. The audio output 2 and input 3 also include a loudspeaker 9 and microphone 10 respectively, to allow the telephone to be used for speech.

The fixed part comprises an interface with the cellular telephone network 11, connected through a DTMF decoder 12 and encoder 13 and a controller interface 14 to a computer 15. The computer 15 comprises a number of servers 16, one of which is allocated to each active mobile unit. The servers 16 have access to a geographical database 17, and a database of standard messages 18. The geographical database 17 is updateable through updating input 19. The database 17 stores the definitions of a number of overlay areas which together form a geographical overlay to the territory to be covered. Examples of overlays are illustrated in Figures 2, 4, 5a, 5b, and 6, to be described in detail later.

The mobile part obtains location information using

the GPS receiver 7 and transmits this information, together with a request for directions to a specified destination, to the fixed part, where a server 16 relates the location information to its geographical database 17 and obtains message information associated with the location from the database 18, and transmits the information back to the mobile part.

The computer 15 may transmit messages in DTMF code, using the encoder 12, or it may generate voice messages which are transmitted through a voice output 20 to the cellular network 11.

DTMF signals are used to transmit the position of the vehicle to the computer 15 which can then offer information and guidance either to the vehicle or to a third party on demand.

In the following discussion, variations on the basic apparatus depicted in Figure 1 will also be described, in which certain elements are modified or replaced.

The system is operated as follows:-

At the start of a journey the driver requests service by activating a pre-dialled control on the telephone 1. This service request is transmitted to the control interface 14 over the telephone network 11. The control interface 14 then allocates a free server 16 to answer the call and interrogate the vehicle GPS receiver 7 to determine its geographical position. The encoder 8 takes the latitude and longitude data and translates the numbers into DTMF tone-pairs, in a manner to be described in more detail below.

The cellular telephone couples this audio signal into its speech input path. This is easy to do with a hands-free vehicle-mounted cellular telephone since the microphone lead is accessible or alternatively, a small transducer can be mounted next to the microphone 10. A DTMF receiver 5 coupled to the loudspeaker 9 (again acoustically or electrically) decodes supervisory data (again in DTMF format) coming back from the server 16 to acknowledge the reception of location messages. If no acknowledgement is received by the DTMF unit then the data message is repeated.

The fixed end of the system comprises a DTMF decoder 12 and encoder 13 coupled to a serial data interface 14 of the server computer 15. This computer, on the one hand, can call the mobile part which will answer automatically and then provide its location using the DTMF signalling system or on the other hand can receive an unsolicited call, which would include the DTMF encoded identity of the mobile unit and would also provide the vehicle location using the DTMF interface 6.

The server 16 then captures the current position of the user, and identifies the overlay area within which that position falls. The server also captures any permanent user-specific information such as the type of vehicle, which may be relevant for the route to be selected e.g. because of height or weight restrictions. The user may encode those requirements which are not permanent, but are specific to the present information request, (in particular his destination) by using the telephone keypad in response to voice prompts. However, in a preferred arrangement the call is presented to a human operator for the capture of this data. This allows the user to obtain assistance in identifying his desired destination to the system, and also allows the driver to speak his requirements, keeping his hands and eyes free for driving.

The operator then remotely programs the in-vehicle interface 6 with system data identifying the vehicle destination, for use in subsequent update processes, and instigates the generation of voice given directions and instructions to the driver by a speech generation subsystem of the computer server 16.

Position fixes may be made at regular intervals, e.g. every two minutes, or every kilometre. Alternatively the fixed part may request the mobile unit to send its next position fix after a specified interval or distance.

As the driver follows the route further instructions can automatically be sent as the driver enters each new overlay area and the driver can be alerted if the route has been left or if any new traffic problems have been detected that will affect the individual driver. The system is arranged such that when the system locates a mobile unit entering an overlay area having a message defined for it, for example the next turn instruction (or an error message if the mobile unit has gone off the selected route), that message is transmitted. The system may also be arranged to transmit messages to users other than the mobile unit in question, for example to monitor the progress of valuable cargoes.

At any time the driver can call the human operator if service requirements change or additional help is needed.

Because a central database is used all vehicle movements can be monitored. Traffic models can be used to optimise traffic flows and reduce journey times. The system can also ensure that it does not itself cause congestion, by limiting the number of vehicles it directs to use the same road at the same time. The control system can use the location data to calculate and record movement vectors from these vehicles.

Using the data collected by this method, it is possible for the central system to derive a digital map of valid routes. The following data could be derived automatically: valid travel lanes; permitted direction(s) of flow; allowable turns; average travel times; trends in travel times according to time of day and other factors.

The system would automatically update the map to show permanent changes (new road inks, changes to one way systems etc.). Temporary lane closures from road works etc. would also be recorded. Manual updating of data would be necessary (for instance to alert the system to a new bypass opening) before the system acquired the information from vehicle flow data, to ensure vehicles are routed over the new road initially. Any approximations in the pre-entered data would automatically be corrected by the system described here.

The system could be further enhanced to include any other information that may be relevant to travellers, by a combination of manual and automated data entry, e.g. location of bus stops, telephone boxes and other street furniture, and proximity to enterprises such as shops, banks or offices.

The variation of transit time trends according to time of day, for each link, could be used to derive a congestion prediction model, as the basis for route guidance. The system may monitor the progress of the mobile units along the routes selected for them, to identify any areas of traffic congestion etc, by comparing actual transit times between predetermined locations. This may be done by the fixed system monitoring the location updates of individual units, or it may be done by the mobile unit, in co-operation with the fixed unit. In this latter case, the fixed part transmits an expected range of transit times within which the mobile is expected to reach a predetermined location. If the mobile unit reaches the location outside this range, it reports the fact to the fixed part. By "reporting by exception" the data processing overhead can be reduced considerably.

However, these systems can become unstable if too many drivers have access to route guidance based on information about current or predicted congestion. To avoid these instabilities route plans are created and updated centrally and passed to individual vehicles. The impact of these vehicles using the suggested routes is then added to the prediction. As more vehicles use the system the prediction produced could become more accurate.

The routes derived can be passed to the vehicles (via a mobile data link, or possibly a short range communications link or other temporary access to a fixed telecommunications network - prior to departure). The vehicle would then operate autonomously, unless the road conditions varied significantly from those predicted.

If the central system detected a problem (from vehicle data or other sources), which had a severe impact on predictions, sufficient to cause a change to advice already given, then the central system could broadcast news of the problem, such that those vehicles affected could automatically call in via a mobile data communications link to receive a new route from its present location to its destination.

If a vehicle system encountered unexpected transit times along its programmed route it would send a report to the central system.

The data flowing though the system will therefore allow it to "learn" more of the road network's characteristic congestion behaviour, e.g. by use of neural net techniques, and to select routes for traffic which avoid using routes at times when they are likely to be congested. In addition, the system can generate digital road maps or other data automatically, based on the position measurements of vehicles using the roads.

A particular advantage of this system is the ability to predict unusual patterns of congestion from the route guidance information requested by the users. Because route guidance is generated centrally, the system can monitor the number of requests for destination information to a given location. By determining the predicted arrival times for each user (which will depend on their starting points, and the time the journey started), a build-up of traffic converging on a particular location at a particular future time (e.g. for a major sporting event) can be detected. Traffic for other destinations, which might have been routed by way of this location, can then be diverted to other routes.

The system described above uses an analogue telecommunications link, in which DTMF codes may be used. For an analogue cellular radio network DTMF is an ideal signalling medium when only short status messages are required to be transmitted. It can survive in the severe signal fading and noise of the mobile environment which frequently precludes the use of fast phase or frequency shift data modulation. Another advantage is the ability to co-exist with speech. For example a DTMF data burst containing vehicle position data could be sent at the start of a call and at intervals during the call. Other simple coded DTMF messages can also be conveyed to indicate emergencies, provide simple driver indications (e.g. illuminated arrows to turn left or right) or trigger synthetic speech generated by another sub-system in the vehicle.

The DTMF coding described above is suitable for an analogue system. In a digital cellular network digitised data can be transmitted over an associated packet data system such as the Short Message Service (SMS) of GSM (Global System for Mobile Communications), or the General Packet Radio Service (GPRS) proposed for GSM.

In the embodiment described above, the speech generation subsystem forms part of the server 16. Alternatively, it can be carried on board the vehicle. In this arrangement the subsystem has various stored speech commands which are controlled from the in-vehicle interface 6 in response to commands transmitted from the fixed part. This arrangement reduces the signalling traffic required over the radio link 11, but increases the complexity of the in-vehicle equipment.

The location-determination system will now be described in greater detail. GPS (Global Positioning System) satellite navigation receivers are now becoming very cheap and are available with a serial data output. These can provide latitude and longitude data to within a tenth of a second of arc (defining position to within 3 metres, which is sufficient to identify which carriageway of a dual carriageway road a user is on).

Satellite positioning systems such as the Global Positioning System (GPS) are prone to small systematic errors, for example as a result of instabilities in the orbits of the satellites. The accuracy of the position measurement may be enhanced by a process known as "Differential GPS" in which a number of fixed reference

points are used, whose positions are determined with great precision e.g. using surveying techniques. GPS is used to obtain a measure of the position of one or more of the fixed reference points. This measure is compared with the known, true location to generate a correction value which can be used to correct the position of the mobile unit as measured by GPS.

The position data received from the satellite positioning system may include some redundant data. If the system is only to operate within a limited area of the globe the most significant digits of the position data are redundant, and need not be transmitted from the mobile unit to the fixed part. For example, any point in Germany can be uniquely defined by the units digits of its latitude and of its longitude, as that country lies entirely between 45 and 55 degrees North, and between 5 and 15 degrees East. It is also possible to define any point in the United Kingdom in this way, although in that case a 10 degree offset in longitude has to be applied to avoid duplication of longitudes East and West of the zero meridian.

For larger territories e.g. a pan-European system, or one covering the USA, this simple method of data reduction is impractical. However, it is nevertheless possible to reduce the data requirements by dynamically defining the territory. After an initialisation step using the full location, the system selects as each new location the closest candidate to the previous one. For example, if the mobile unit was last reported at 99 degrees W and the units digit of the longitude is now 0, the user is taken to be at 100 degrees W rather than, for example, 90 degrees or 110 degrees.

If location updates take place sufficiently frequently that the user's position cannot have changed by more than half a degree, the units digit of degrees may also be dispensed with, and the location given only in minutes and seconds of arc. The more frequent the updates, the more digits can be dispensed with.

An alternative method of obtaining the coarse position location is interrogation of the cellular radio system's operating system to identify the cell in which the user is currently located. Cell sizes can be up to about 40km across (although they are often much smaller, so identifying the cell can identify the user's location to within 40km, which identifies latitude to better than half a degree. (1 degree of latitude = 111km). The separation of lines of longitude varies with the cosine of the latitude but even at the Arctic Circle (66 degrees North) a 40km resolution will identify longitude to the nearest whole degree (1 degree of longitude = 111 km (cos latitude) = approximately 45km at 66 degrees North).

By left-truncating the position data by omitting the degrees digits a basic position message would therefore consist of 10 decimal digits (minutes, seconds, and tenths of seconds). Altitude data giving altitude in metres would require a further four digits, since all points on the Earth's surface lie within a range of 10,000 metres, but this data can also be left-truncated, as it is unlikely that any multi-level road system would exceed 100 metres in height (or if it did, that a GPS system would work effectively for any receiver on the lower levels). This gives a total of twelve digits, which can be transmitted by DTMF in less than 2 seconds.

If the data is left-truncated as described above, the "coarse" data is added by the interface controller 14 by reference to the previous position or to the cellular radio operating system.

When the computer 15 receives a location message, it stores the location and then searches its database for an overlay area within which that position lies. The overlay areas are defined in the database by co-ordinates of latitude and longitude and have associated attributes which define messages which can be passed to mobile subscribers within the overlay area defined. In some instances height (altitude) information, also available using satellite positioning systems, may be used, for example to distinguish between levels in a multi-level highway intersection. When a DTMF location message has co-ordinates which fall inside an overlay area having an associated message, the message is then transmitted to the mobile part as a computer synthesised speech message, a DTMF coded message (to activate other subsystems) or as a high speed conventional data message.

If the mobile unit fell within the same overlay area at the previous location update, and the message associated with that overlay area is unchanged, the transmission of the message may be suspended.

The frequency at which location updates are requested by the system may be tailored to the size and nature of the current overlay area. For example, an intricate road layout may comprise a large number of small overlay areas, requiring frequent location updates to ensure that a user does not miss an instruction by passing through its associated area between two updates. However, a long stretch of road without junctions may be covered by a single overlay area, so less frequent updates are appropriate. The speed with which a vehicle is likely to be moving, which will differ between urban, rural, and motorway environments may also be used as a factor in determining when the next location update should be requested.

As suggested above, there may be circumstances when a satellite positioning system may be unusable, for example in tunnels or built-up areas where a line-of-sight view of the satellites may be impossible to obtain. Alternative arrangements for identifying and updating the mobile part's location which do not rely on a satellite receiver may be used, either on their own, or to interpolate between points where a satellite system can be used. In one variant, a navigation system based on dead-reckoning may be used. In such systems the user identifies his initial location and the on-board system measures the system's movement e.g. by magnetic bearing measurements, distance counters, and inertial navigation means such as gyrocompasses and acceler-

ometers. Such systems are self-contained, but require knowledge of the starting point. This may be obtained, for example from a satellite positioning system.

In another variant, a method of location may be used which relies on the propagation characteristics of the cellular radio system used for communication with the central control station. Examples of such systems are disclosed in German Patent specifications DE3825661 (Licentia Patent Verwaltungs) and DE 3516357 (Bosch), United States Patent 4210913 (Newhouse), European Patent specification EP0320913 (Nokia), and International Patent applications WO92/13284 (Song) and WO 88/01061 (Ventana). By comparison of signal strength or other characteristics of several cellular base stations, a position fix can be determined. In this arrangement the location measurement may be made directly by the fixed system. This allows the mobile part of the system to be embodied by a conventional cellular telephone, with inputs being provided by speech, or by DTMF tones generated by the keypad, and instructions to the user being transmitted by voice commands.

Examples of the kind of navigation information which may be stored in the database 17 will now be discussed, with reference to Figures 2 to 6. Briefly, Figure 2 shows a junction J having four approach roads 21,22,23, 24; each having associated with it an overlay area 21a, 22a, 23a, 24a respectively. In this figure, and all other figures illustrating road layouts, the roads are shown arranged for left-hand running, as used for example in the UK, Japan, Australia etc. Figure 3 shows part of a road network surrounding the junction J, including towns A, B, C, and a motorway M. Each of the roads 21, 22, 23, 24 has an associated destination zone 21z etc. Figure 4 shows a complex grade-separated junction interlinking four roads N, S, E, W. The junction has superimposed on it an overlay having twelve overlay areas, Na, Ni, Nd, Sa, Si, Sd, Ea, Ei, Ed, Wa, Wi, Wd. Figure 5a shows a small region having a main road 33 and a side road 30. The main road 33 has two associated overlay areas 31, 32. Figure 5b is similar to Figure 5a, but an obstruction X is present on the main road 33, and the overlay area 32 has been subdivided into two overlay areas 32a, 32b, separated by the obstruction. Figure 6 shows an overlay comprising ten overlay areas 40 - 49 superimposed on a cellular radio coverage region comprising five cells 50 - 54.

In greater detail, the road junction J (Figure 2) has four approach roads 21, 22, 23, 24. On each road, at the approach to the junction, an overlay area (21a, 22a, 23a, 24a) is defined. These overlay areas have directional information associated with them, giving turn instructions or other navigational information. As shown in Figure 3, the entire territory covered by the navigation system can be divided into four zones 21z, 22z, 23z, 24z, each comprising the set of all locations for which the corresponding road 21, 22, 23, 24 should be taken from the junction J. In this particular example, road 24 leads directly into town A and is only used for local destinations (zone 24z), road 23 leads to town B (zone 23z), road 22 leads to town D (zone 22z) and road 21 leads to the motorway M, for all other destinations including town C and part of town A. These zones are defined differently for each junction: for example at junction J' different directions are appropriate for towns A and C, so these towns fall in different zones with respect to the overlay areas at that junction. The zones may even be defined differently for different overlay areas at the same junction. For example, if U-turns are not possible at the junction J, any traffic approaching the junction J by road 22 and requiring town D (perhaps as the result of a previous error, or a change of plan) must be routed by way of roads 21, M, and 25. Thus, for overlay area 22a there are only three zones: 24z, 23z and the combined 21z/22z, corresponding to the three permitted exits 21, 23, 24.

The zones may be re-defined according to circumstances. For example, when the motorway M is congested, the best route from junction J to town C may be by way of town B. In such circumstances, zones 21z and 23z are redefined so that town C now falls within zone 23z. It should be noted, however, that the total number of zones remains the number of exit routes from the relevant overlay area.

The overlay areas 21a, 22a, 23a, and 24a should be large enough to ensure that any vehicle approaching the junction gets at least one location update whilst within the relevant overlay area, and is thus sent the relevant turn instruction. As shown in Figure 2, these overlay areas are discrete, and may be considered equivalent to the coverage areas of the beacons of the prior art system discussed above. They may, however, be made contiguous, as shown in Figures 4, 5a, 5b and 6.

Figure 4 shows a more complex, grade-separated junction, in which there are twelve overlay areas. Each road N, E, S, W intersecting at the junction has a corresponding approach overlay area Na, Ea, Sa, Wa, (Wa shown shaded), and a depart overlay area Nd, Ed, Sd, Wd (Ed shown shaded). There are also four intermediate overlay areas Ni, Ei, Si, Wi (Si shown shaded). In the vicinity of the flyover F height (altitude) information obtainable from the GPS system can be used to determine which level, and therefore which overlay area, the user is currently in.

The approach and intermediate overlay areas each end at a decision point P1 to P8. In the database 17 each overlay area has direction information associated with it, providing instructions as to which fork to take at the associated decision point. For example, the direction information associated with zone Si instructs users for destinations served by road N to go straight on at point P1, and users for destinations served by roads E, S, and W to turn left. It will be seen that traffic using the intersection will pass through one approach overlay area, one departure overlay area, and may also pass

through one or more intermediate overlay areas. There may also be information associated with the departure overlay areas Nd, Sd, Ed, Wd, for example warning of hazards ahead. The departure overlay areas may be continuous with approach overlay areas for the next junction in each direction.

As a user approaches the junction on road S, a location update identifies the user equipment as being within overlay area Sa. If the co-ordinates of the user's destination are within the zone served by road W, the user is sent an instruction to turn left at point P2. If the user obeys this instruction, he will enter overlay area Wd and on the next location update he will be sent information relevant to that overlay area (if any).

If the co-ordinates of the user's destination are within the zone served by road N, the user in overlay area Sa is instead sent an instruction to continue straight on at point P2. If the user obeys this instruction, he will enter overlay area Si.

For a user in overlay area Si, if the co-ordinates of the user's destination are within the zone served by road N the user is sent an instruction to go straight on at point P1. On obeying this instruction, he will enter the overlay area Nd and on the next location update he will be sent information relevant to that overlay area (if any).

If the co-ordinates of the destination of a user in overlay area Si are in the zone served by roads E, S, or W, the user will be sent an instruction to turn left at point P1. On obeying this instruction, he will enter overlay area Wi.

Similar information is associated with the other overlay areas. By being given appropriate instructions as the user negotiates a succession of junctions (decision points), the user can be directed to any destination. It should be noted that all users who are to be directed to the same exit from the junction are given the same instruction, whatever their ultimate destination.

Figures 5a and 5b illustrate the reconfiguration of the overlay areas to meet changing circumstances. Initially (Figure 5a) an overlay area 31 is defined for the approach to a junction between a major road 33 and a side road 30, and a second overlay area 32 is defined for that part of the major road 33 beyond the junction. Information associated with the overlay area 31 includes turn information to instruct traffic for the zone served by the side road 30 to turn off. Information may also be associated with the overlay area 32.

In figure 5b the major road 33 has been blocked at a point X. In order to accommodate this, the overlay area 32 has been subdivided into two overlay areas 32a, 32b. The information (if any) associated with overlay area 32b is the same as that previously associated with overlay area 32. Traffic in overlay area 32a is given new information warning it of the hazard ahead. The information associated with the overlay area 31 is modified, so that all traffic is now instructed to turn off onto the side road 30. (Effectively this means that the destination zones associated with the overlay area 31 are merged into one)

Figure 6 shows how the overlay areas may be defined for a road network. In this example there is an overlay area 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 corresponding to each side of each section of road. Information appropriate to each direction of travel on each section is therefore available to users throughout the relevant section. Superimposed on this overlay there is a cellular radio network, five cells of which (50, 51, 52, 53, 54) are shown. The position of the user, as determined for example by a satellite positioning system, determines which overlay area is appropriate to the user. The information is transmitted to the service control centre by means of the cellular radio network. Handovers between cellular base stations occur in conventional manner at cell boundaries. These handovers are, however, unrelated to the boundaries between the overlay areas 40 - 49

Although the described embodiment relates to the provision of route guidance information, other locality-dependant information may be provided as well, or instead, such as information about local facilities, tourist attractions, weather forecasts, public transport information, etc. The term "guidance information", as used in this specification, embraces any such information.

Claims

1. A navigation information system for providing information to one or more mobile users dependent on their locations, the system comprising:

   means (14, 15, 17) for determining the location of a mobile unit requesting guidance information,
   means (18, 15, 14) for generating information for guidance of the user of the mobile unit according to the present location of the mobile unit,
   and a communications system (11, 13) for transmitting the guidance information so generated to the mobile unit,

   wherein the navigation information system has means (12,13) for communicating with the mobile unit using dual tone multi-frequency (DTMF) signals.

2. A system as claimed in Claim 1, having means (12) for receiving and decoding information in DTMF format relating to the location of the mobile unit.

3. A system as claimed in claim 2, having means (14) for determining a coarse position of the mobile unit, means (12) for receiving fine position information from the mobile unit, and means for combining the coarse and fine position data to define a unique poistion measure.

4. A system as claimed in claim 3, having means for recording a previous known position, and means for retrieving said position data as the coarse position data.

5. A system as claimed in claim 3 or 4, having means for determining the approximate location of the mobile unit by reference to its interaction with a cellular radio system.

6. A system according to any preceding claim, having means for generating coded DTMF messages for transmission to the mobile units.

7. A mobile unit for a navigation information system, comprising means (7) for identifying the present position of the mobile unit, means (8, 10, 1) for transmitting to a co-opertaing fixed part, over a communications link (11), information relating to the present location of the mobile unit, and guidance instruction means (6) controllable by guidance instruction information received over the communications link, whereby guidance instructions related to the present location can be communicated to a user by means of the guidance instruction means,

characterised in that the mobile unit has means (5,8) for communicating with the co-operating fixed part using dual tone multi-frequency (DTMF) signals.

8. A mobile unit as claimed in Claim 7, having means (8) for encoding and transmitting information in DTMF format relating to the location of the mobile unit.

9. A mobile unit according to claim 8, having a satellite position finding means (7) for determining the position of the mobile unit.

10. A mobile unit according to claim 9 comprising means for transmitting the position data, truncated by the omission of one or more of the most significant digits.

11. A mobile unit according to claim 9 or 10, comprising means for adding a predetermined offset to the position data provided by the satellite system.

12. A mobile unit according to claim 8, having an inertial means for determining the position of the mobile unit.

13. A mobile unit according to any of claims 7 to 12, having means (5) for receiving and decoding DTMF messages transmitted from the fixed part, and means (6) controllable in response to said DTMF messages.

14. A mobile unit acccording to claim 13, wherein the means (6) controllable by the received DTMF messages is a voice synthesiser.

15. A method of providing navigation guidance information to mobile units of a mobile radio system, the information being dependent on the locations of the mobile units, the method comprising the steps of:

- transmitting, from a mobile unit (1) to a fixed part (11-20), a request for navigation guidance;
- determining the location of the mobile unit;
- generating guidance information on the basis of the location information and navigation data stored in the fixed part; and
- transmitting the guidance information from the fixed part to the mobile unit;

characterised in that at least part of the data transmitted to and/or from the mobile unit is in the form of DTMF signals.

16. A method according to claim 15, wherein DTMF signals are used to transmit the position of the mobile unit to the fixed part

17. A method according to claim 16, wherein the mobile unit determines its position using a satellite position finding means (7).

18. A method according to claim 17 wherein the mobile unit transmits the position data to the fixed part truncated by the omission of one or more of the most significant digits, the fixed part reinstating the truncated digits according to coarse position data determined by the fixed part.

19. A method according to claim 18, wherein the fixed part retrieves a previously recorded known position of the mobile unit for use as the coarse position data.

20. A method according to claim 18 or 19 wherein the coarse position data is determined by reference to the interaction of the mobile unit with the mobile radio system.

21. A method according ing to claim 18 19 or 20, wherein a predetermined offset is applied to the position data for transmission from the mobile unit to the fixed part, and the fixed part removes the predetermined offset to determine the true position of the mobile unit.

22. A method according to claim 16, wherein the mobile unit determines its position by inertial means.

23. A method according to any of claims 15 to 22 wherein DTMF coded messages are transmitted

from the fixed part to the mobile part.

24. A method according to claim 23, wherein the messages are arranged to activate a subsystem of the mobile part.

25. A method according to claim 24, wherein the subsystem is a voice synthesiser.

Fig.1.

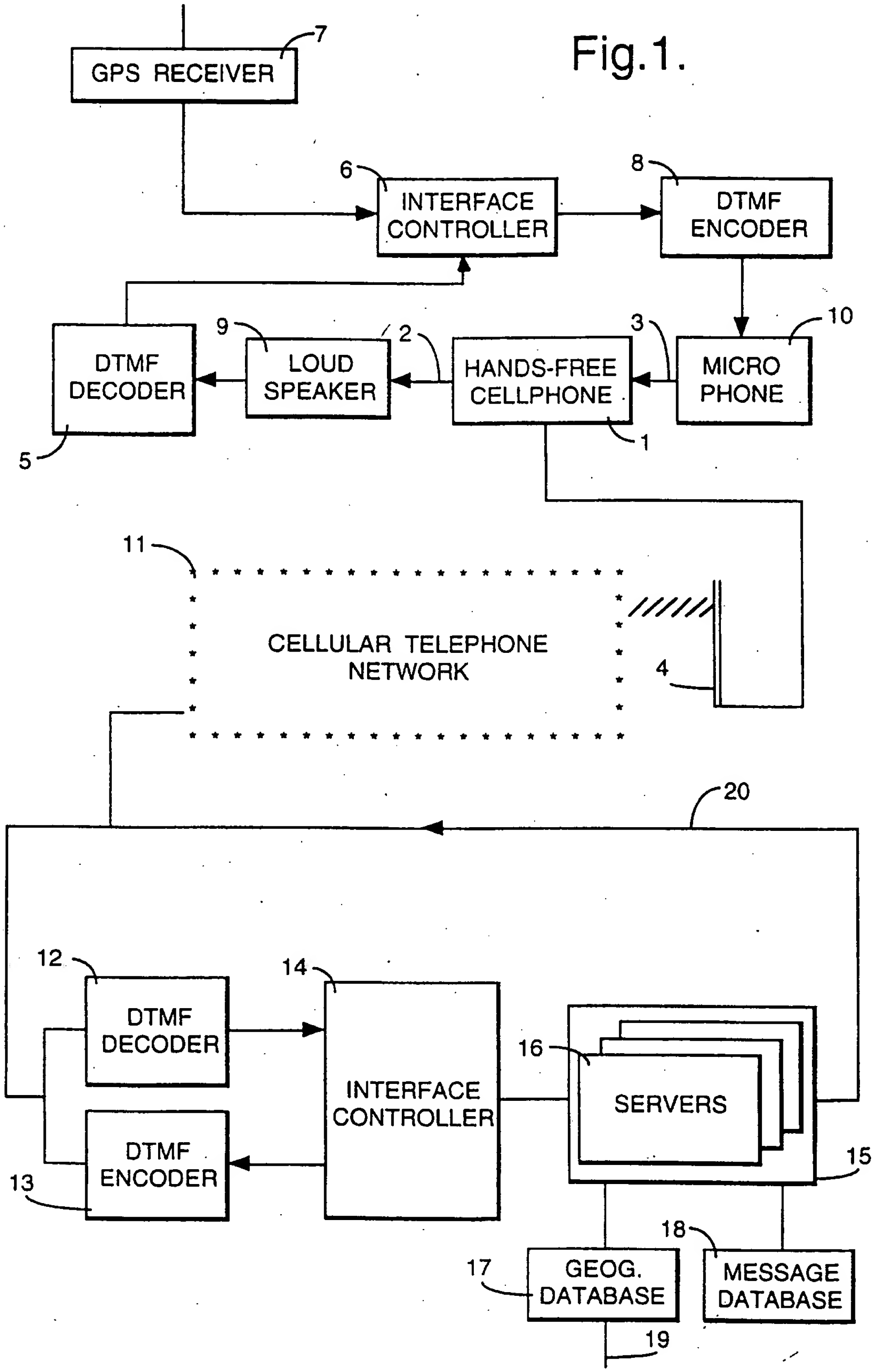

GPS RECEIVER
7
6
INTERFACE CONTROLLER
8
DTMF ENCODER
9
LOUD SPEAKER
2
HANDS-FREE CELLPHONE
3
MICRO PHONE
10
DTMF DECODER
5
1
11
CELLULAR TELEPHONE NETWORK
4
20
12
DTMF DECODER
14
INTERFACE CONTROLLER
16
SERVERS
15
DTMF ENCODER
13
18
17
GEOG. DATABASE
MESSAGE DATABASE
19

Fig.2.

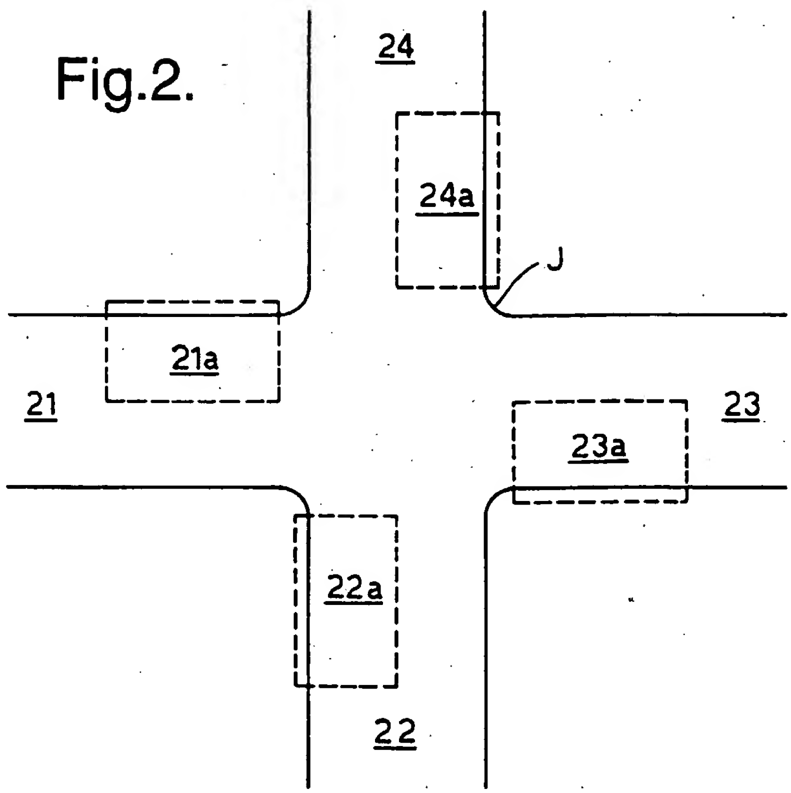

24
24a
J
21a
21
23
23a
22a
22

Fig.3.

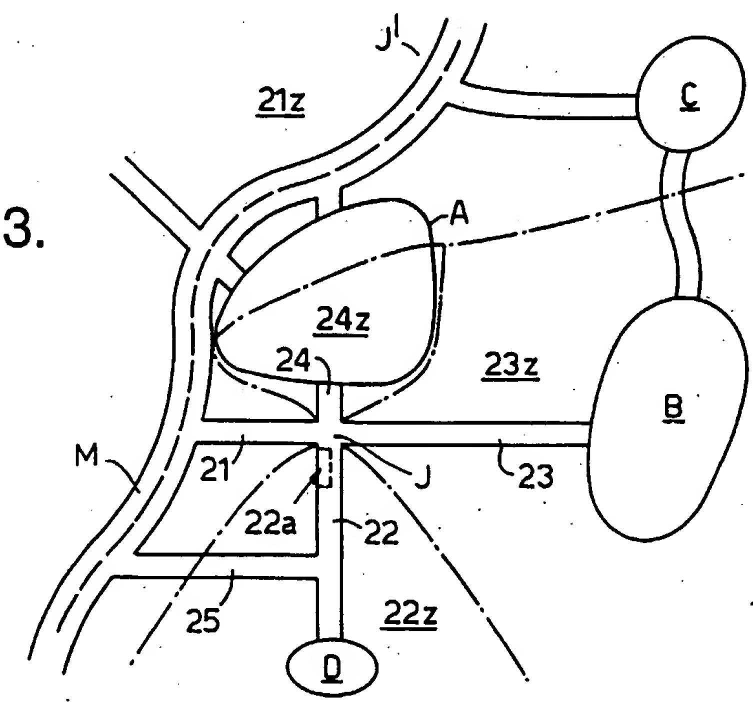

J'
21z
C
A
24z
24
23z
B
M
21
J
23
22a
22
25
22z
D

Fig.4.

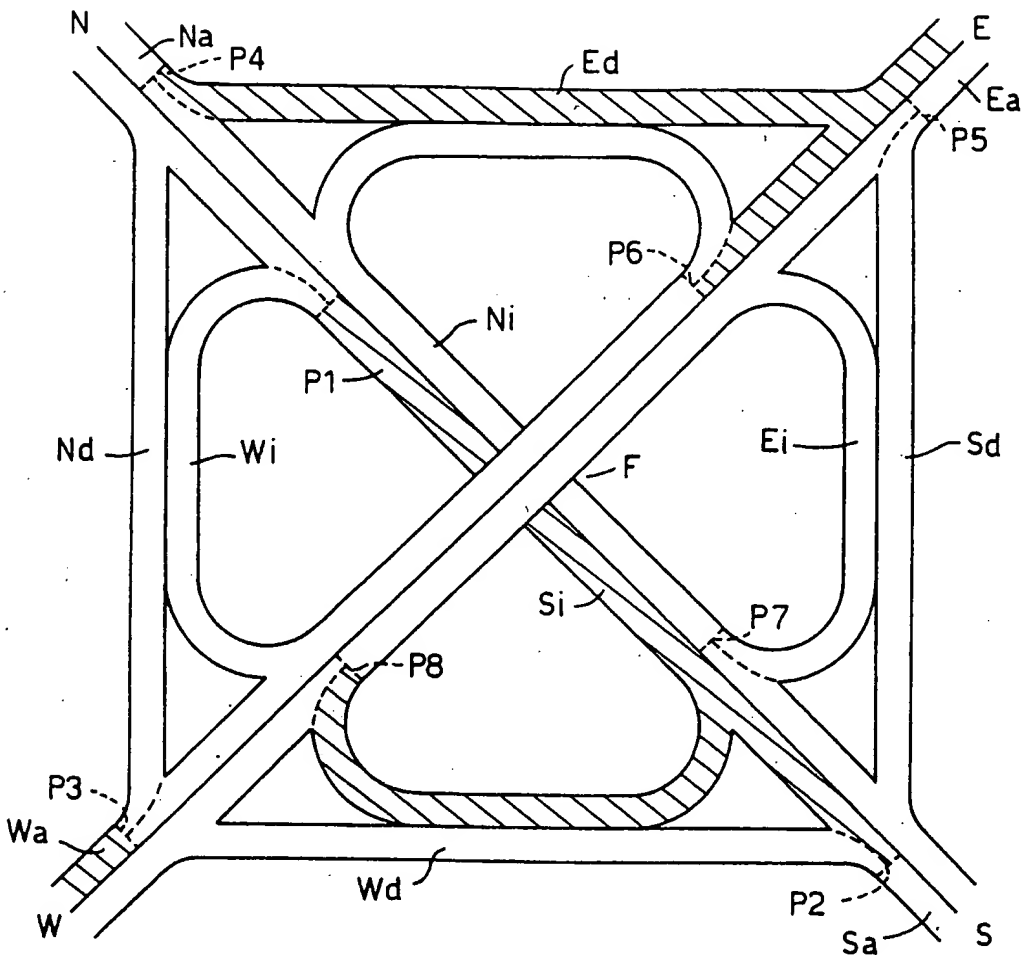


Fig.5a.

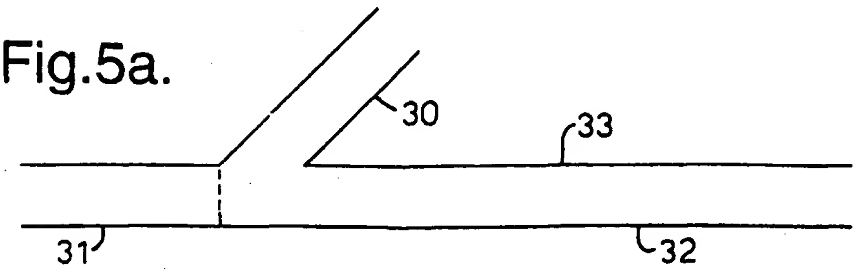


Fig.5b.

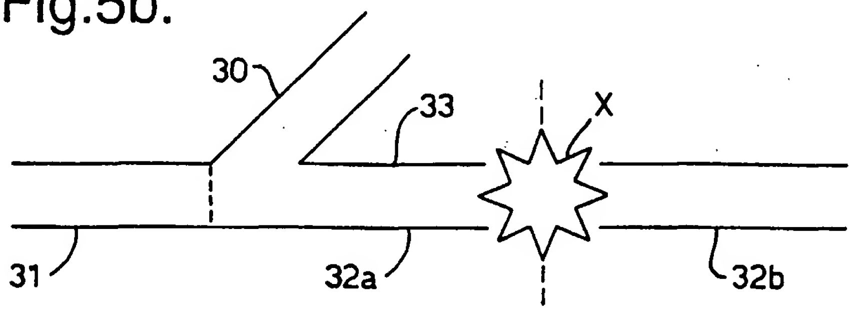
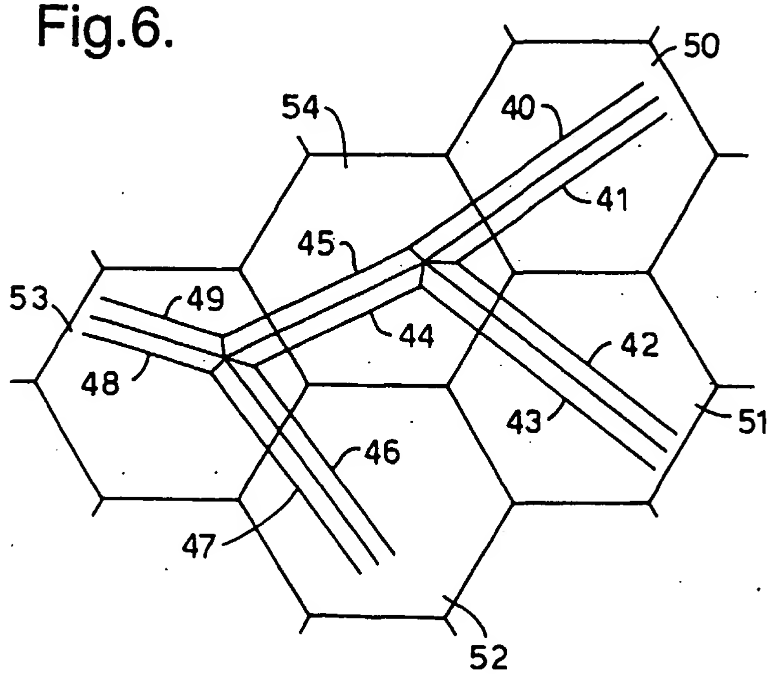


Fig.6.



European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 97 20 3380

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Y	EP 0 123 562 A (BRITISH TELECOMM) 31 October 1984 * page 2, line 32 - page 4, line 13 * * page 4, line 22 - page 5, line 15 * * page 6, line 20 - page 7, line 21 * * page 9, line 1 - line 5; figures 1,2 * ---	1-3,6-9, 12-17, 22-25	G01S5/14 H04Q7/38 G08G1/127
Y	US 5 081 667 A (DRORI ZE EV ET AL) 14 January 1992 * column 4, line 12 - line 48 * * column 5, line 9 - line 17 * * column 11, line 43 - line 59; figure 1 * ---	1-3,6-9, 12-17, 22-25	
A	EP 0 345 818 A (OKI ELECTRIC IND CO LTD) 13 December 1989 * the whole document * ---	1,3,5,7, 14,15,25	
A	US 5 208 756 A (SONG HAN L) 4 May 1993 * the whole document * -----	1,6,7, 13,15, 23,24	TECHNICAL FIELDS SEARCHED (Int.Cl.6) G01S H04Q G08G

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	11 February 1998	Haffner, R

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

L1: Entry 33 of 35 File: EPAB Apr 22, 1998

PUB-NO: EP000837341A1
DOCUMENT-IDENTIFIER: EP 837341 A1
TITLE: Navigation information system

PUBN-DATE: April 22, 1998

INVENTOR-INFORMATION:

NAME	COUNTRY
MANNINGS, ROBIN THOMAS	GB
WALL, NIGEL DAVID CHARLES	GB

ASSIGNEE-INFORMATION:

NAME	COUNTRY
BRITISH TELECOMM	GB

APPL-NO: EP97203380
APPL-DATE: September 1, 1995

PRIORITY-DATA: EP95930610A (September 1, 1995), GB09417600A (September 1, 1994)

INT-CL (IPC): G01S 5/14; H04Q 7/38; G08G 1/127
EUR-CL (EPC): G01C021/34; G01S005/00, G01S005/14 , G08G001/0967 , G08G001/0968 , G08G001/127

ABSTRACT:

CHG DATE=19990617 STATUS=O> A navigation information system comprises a communications system having a fixed part (11 to 20) and at least one mobile part (1 to 10), the fixed part including a data storage and processing means 15 for identifying the location of a mobile unit, generating guidance information appropriate to that location and transmitting it to the mobile unit. The data exchange is by means of DTMF codes. By locating most of the complexity with the service provider, in particular the navigation computer 15 and geographical database 17, the system can be readily updated and the capital cost of the in-vehicle system, which in its simplest form may be a standard cellular telephone 1, can be minimised. The user makes a request for guidance information, and the system, having determined the user's present location, then transmits instructions to the user. The user's present location can be determined by means such as a Satellite Positioning System 7.

PCT

WORLD INTELLECTUAL
Internati

INTERNATIONAL APPLICATION PUBLISHED UN

WO 9607110A1

(51) International Patent Classification 6 : G01S 5/14, 5/00, H04Q 7/38, G08G 1/127 | A1

(11) International Publication Number: WO 96/07110

(43) International Publication Date: 7 March 1996 (07.03.96)

(21) International Application Number: PCT/GB95/02065

(22) International Filing Date: 1 September 1995 (01.09.95)

(30) Priority Data:
9417600.5 1 September 1994 (01.09.94) GB

(71) Applicant (*for all designated States except US*): BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY [GB/GB]; 81 Newgate Street, London EC1A 7AJ (GB).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): MANNINGS, Robin, Thomas [GB/GB]; 12 Mayfields, Martlesham Heath, Ipswich, Suffolk IP5 7TU (GB). WALL, Nigel, David, Charles [GB/GB]; 9 North Close, Ipswich, Suffolk IP4 2TL (GB).

(74) Agent: LIDBETTER, Timothy, Guy, Edwin; BT Group Legal Services, Intellectual Property Dept., 13th floor, 151 Gower Street, London WC1E 6BA (GB).

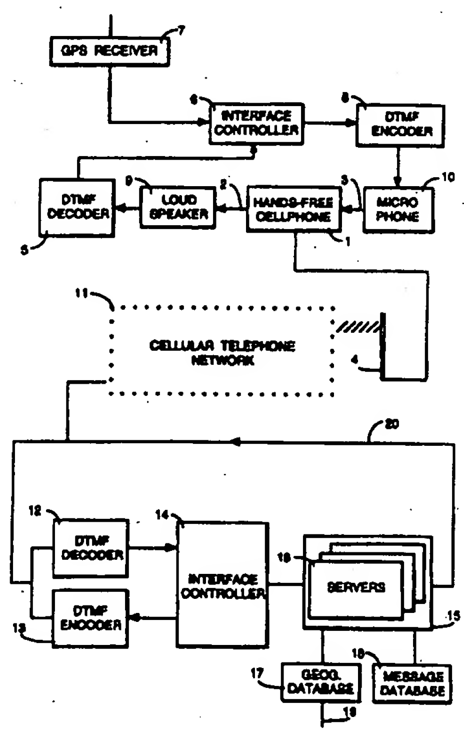
(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: NAVIGATION INFORMATION SYSTEM

(57) Abstract

A navigation information system comprises a communications system having a fixed part (11 to 20) and at least one mobile part (1 to 10), the fixed part including a data storage and processing means (15) for identifying the location of a mobile unit, generating guidance information appropriate to that location and transmitting it to the mobile unit. By locating most of the complexity with the service provider, in particular the navigation computer (15) and geographical database (17), the system can be readily updated and the capital cost of the in-vehicle system, which in its simplest form may be a standard cellular telephone (1), can be minimised. The user makes a request for guidance information, and the system, having determined the user's present location, then transmits instructions to the user. The user's present location can be determined by means such as a Satellite Positioning System (7).



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

NAVIGATION INFORMATION SYSTEM

This invention relates to navigation information systems. It is particularly suitable for use in providing users of road vehicles with route guidance, but other applications are possible and are discussed below.

Navigation of a vehicle through an unfamiliar complex road network is a difficult task. Large amounts of fuel and time are wasted as a result of drivers getting lost or using an inefficient route. Accidents can also be caused by drivers attempting to read maps or complex road signs and losing concentration on the road ahead. Moreover, a driver may choose an inefficient route as a result of using an out-of-date map.

An additional problem can occur even if a driver knows a route to his or her destination. That route may be congested or blocked as a result of accidents or maintenance work, so that an alternative route would be more efficient.

Several proposals have been made for navigation guidance systems. In some such proposals a vehicle-borne system has a navigation computer and a geographical information system which is essentially a digitised map stored on a CD-ROM. The system gives the driver information and guidance by screen and/or speech display. These systems would be very expensive. Each vehicle requires a navigation computer and geographical information system. The cost of the complex vehicle-borne equipment involved is estimated to be in the region of £1000. The system is complex to operate, and could only be safely operated by the driver whilst the vehicle is stationary. The geographical information system would require periodic updating, which requires new disks to be distributed to subscribers from time to time.

In some proposed systems of this type real-time data would be broadcast over a radio network to update fixed information held on the geographical information system. Even so, the geographical information system would only be accurate up to its last update. Moreover, a broadcast channel needs to be allocated for the updating service.

It has also been proposed that the guidance service provider collects statistical traffic flow data from which traffic congestion predictions can be made which are fed into the real-time data to be broadcast. The traffic flow data may be

collected using roadside sensors, or they may be collected by monitoring the operation of the mobile user equipment. The latter approach can only collect data relating to users of the system, but has a lower capital cost.

In an alternative approach a system of short-range roadside beacons is used to transmit guidance information to passing vehicles equipped with simple transceivers. The beacons transmit information to suitably equipped passing vehicles to give turn instructions appropriate to their chosen destinations. For each beacon the territory to be covered is divided into as many zones as there are exits from the junction the beacon relates to. The zone in which the user's chosen destination falls is determined, and instructions are given appropriate to that zone. At any given beacon all vehicles whose destinations are in the same zone get the same instruction. The definitions of the zones are dependant on the location of the beacons, and each zone comprises the set of destinations which should be reached from the beacon by taking the direction associated with that zone.

Each beacon only gives instructions for reaching the next beacon along the route to the vehicle's destination. For two vehicles starting from the same point for different destinations for which the routes are initially coincident, the beacons along the coincident section of route will each give both users the same instructions, because for those beacons both users are travelling to the same zone. Only for the beacon at the point of divergence are the two users' destinations in different zones, and therefore different instructions are given.

The beacons' programming may be modified from time to time by control signals from a central control station, in a way analogous to remotely controlled adjustable signposts, but in its interactions with the user equipment the beacon is autonomous, identifying which of its zones the user's desired destination is in, and transmitting the appropriate "turn" information to get it to the next beacon on the way. The beacon has no knowledge of the rest of the route.

Each beacon has a detailed map of a small local area (the boundaries of which are, in fact, the adjacent beacons), and if the destination is in this area the beacon gives full information of the route to the destination. The system can therefore provide a user with directions to a destination defined more precisely than the beacon spacing. However, at the beginning of a journey, a user cannot use the system until he encounters a beacon.

This proposed system allows instant updating of the guidance instructions from a central control, and simpler in-vehicle equipment, but requires vast capital expenditure in roadside beacons.

A problem encountered with both the proposed systems described above is that it is difficult for them to provide alternative routings in response to congestion, either current or future, without the risk of creating worse problems on the alternative routes. Although predictions of regularly occurring congestion peaks are relatively simple to programme into the guidance information, and, at least in the beacon system, real-time updates on road congestion can also be fed to the programming of the beacons, the control system does not have any information of vehicle movements from which to predict future congestion. In any case, if the system is in use by a significant fraction of the vehicles, the system will tend to produce congestion on the diversionary routes.

According to a first aspect of the invention, there is provided a navigation information system for providing information to a mobile user dependant on the location of the mobile user, the system comprising a mobile communications system having a fixed part and one or more mobile part for communicating with the fixed part, the one or more mobile part including means for transmitting to the fixed part a request for guidance information and for receiving guidance information from the fixed part, and the fixed part including:

means for determining the location of a mobile part requesting guidance information,

means for generating guidance information according to the location of the mobile part, and

means for transmitting the guidance information so generated to the mobile part,

whereby information dependant on the location of the mobile unit can be transmitted to the mobile unit.

According to a second aspect of the invention, there is provided a navigation information system for providing information to one or more mobile user dependant on the location of the one or more mobile user, the system comprising:

means for determining the location of a mobile unit requesting guidance information,

means for generating guidance information according to the location of the mobile unit,

and a communications system for transmitting the guidance information so generated to the mobile unit,

whereby information dependant on the location of the mobile unit can be transmitted to the mobile unit.

According to a third aspect of the invention there is provided a mobile unit for a navigation information system, comprising means for receiving guidance instruction information over a communications link, and guidance instruction means controllable by the guidance instruction information received over the communications link, whereby guidance can be communicated to the user by means of the guidance instruction means.

According to a fourth aspect of the invention, there is provided a method of providing navigation information to mobile units of a mobile radio system dependant on the locations of the mobile units comprising the steps of storing navigation data in a fixed part, transmitting a request for navigation guidance from a mobile unit to the fixed part, determining the location of the mobile unit, generating guidance information on the basis of the stored data, location information and the request, and transmitting the guidance information from the fixed part to the mobile unit, whereby information relevant to the location of the mobile unit is transmitted to the mobile unit.

This invention has advantages over both the prior art systems discussed above. Considerable improvements can be made over the prior on-board navigation system proposals by putting the intelligence in the fixed part of the system. Firstly, there is no need to distribute maps or updates to subscribers because the data is held centrally. New roads can be added to the system at the instant they are opened. Total capital expenditure is minimised since all users share the same database. Moreover, the computing resources are used more efficiently, because

an in-vehicle system spends most of its time inactive but a centralised system can be time-shared.

Moreover, in contrast to the prior art roadside beacon system, the invention can be implemented with little deployment of equipment in the field, thereby offering considerable economies in capital cost and maintenance, and allowing rapid installation and modification of the system to meet changing requirements.

Preferably the system includes means for determining the location of the mobile part in relation to a geographical overlay comprising a plurality of overlay areas, and means for transmitting information associated with an overlay area which includes the location of the mobile part, whereby a mobile part within that overlay area receives information associated with that overlay area. This allows information associated with a particular overlay area to be transmitted to any mobile units in that overlay area. The system may also comprise means for determining when a mobile part enters a predetermined overlay area, and means for transmitting a message, to a user other than the said mobile part, in response to the said mobile part entering the predetermined overlay area. For example, one overlay area may cover part of a road approaching a junction, and the message may be the appropriate instruction to the driver, as he approaches the junction, as to which way he should turn. Each individual overlay area therefore gives navigation instructions specific to that overlay area. The overlay areas may overlap, and may be of any size down to the practical minimum of the resolution of the location determination process. Large overlay areas are suitable for transmitting general information, whilst smaller areas can be used to target information to users in very precise locations, such as individual elements of a complicated road layout. The overlay areas may be delimited in two or three dimensions.

An advantage of this preferred arrangement over the fixed beacon systems is that the geographical overlay can be readily modified. Advantageously, the system includes means for storing a digital representation of the geographical overlay, and means for modifying the stored representation such that the configurations of the overlay areas may be selected to meet changing requirements. The overlay areas can be readily combined or subdivided, or their

boundaries otherwise altered to meet changing circumstances without any modification to the hardware, simply by reconfiguring the geographical overlay defined in the central database. Moreover, unlike the prior art beacon system discussed above, there is no major cost in street furniture and supporting infrastructure, because existing cellular mobile communications systems may be used to transmit the instructions from a central database. If the driver enters an overlay area which is not on the route chosen by the system, an error message can be transmitted. Such messages may be transmitted to a user other than the mobile unit, for instance in order to monitor the whereabouts of valuable cargoes or of personnel working away from a base.

The geographical overlay may also be used to operate an access-control system, for example for site security or for levying tolls. In this arrangement, if a user enters an overlay area for which he does not have permission, an alert signal can be sent to a system controller, or to security staff on site who can intercept the interloper. Means may be provided (either in a fixed location or with the mobile user) to store a value associated with the mobile unit, and means arranged to modify the stored value in response to the messages transmitted in accordance with the location of the mobile unit, either to increment the value e.g. for subsequent billing, or to decrement the value e.g. in a prepaid stored-value device.

The fixed part may include means for storing map information or other data for use in providing information, herein referred to as guidance data, means for updating the stored guidance data, means for identifying mobile parts to which the updated data are applicable, and means for transmitting such data over the communications system to the mobile parts so identified. This allows information about changing traffic situations to be transmitted to all users who will be affected, without needing to broadcast the details to other users as would be the case with those prior art systems where updating is possible.

Although the information transmitted to the user is specific to the location, information about the user can be processed centrally. This allows short-term traffic predictions to be made. The guidance data transmitted to the mobile units can therefore be based on the position measurements of a plurality of the mobile parts. If the mobile parts are vehicles, these position measurements will identify the locations of roads, and an indication of their traffic density. As new roads are

built or routes are diverted, traffic will move to the new routes. Measuring the position of the traffic will therefore result in the data being updated automatically. To reduce the volume of information transmitted, the fixed part may comprise means for transmitting to the mobile part an expected range of movement information and for receiving from the mobile part movement measurements outside the expected range, and the mobile part comprising means for measuring location and time to derive movement information, means to compare the movement information with the expected range received from a fixed part of the system, and means to automatically report to the fixed system movement measurements outside the expected range. In this way only exceptional traffic conditions are reported.

The fixed part may include means for generating and maintaining guidance data based on vehicle movement data derived from time information and position measurements of a plurality of the mobile parts and/or estimations of future locations of the mobile parts based on the guidance information previously transmitted to the mobile parts. Estimations of future locations of the mobile parts based on the guidance information previously transmitted to the mobile parts can be used to make estimates of future traffic situations.

The data stored in the data storage means may be updated, for example in response to changing traffic conditions, accidents, or highway maintenance. The system may include means for identifying the mobile units to which the updated data are applicable, and transmitting amended instructions over the communications system to said mobile parts. With knowledge of the journeys being planned by a large number of users, a better prediction of demand for particular roads (and hence of congestion on those roads) can be built up. This can be more stable than existing autonomous route-planning systems because the navigation system can take account of the journeys planned for other users.

Advantageously the invention can be implemented using a public cellular radio data service on an individual dial-up basis, providing a simple mechanism for billing and avoiding the need for a separate radio transmission system.

The means for determining the location of the mobile part may comprise means to interrogate a location-identifying means forming part of the mobile part operating for example by means of dead reckoning from a known start point, using

an inertial navigation system or distance and direction measuring devices such as a compass and an odometer. Alternatively, the means for locating position may include means for identifying the location of the mobile part in relation to elements of the fixed part of the communications system. The location of the mobile part may be determined by a radio location system associated with the cellular radio system. In another alternative arrangement, a satellite navigation system may be used. In one preferred arrangement the fixed part has means to determine the approximate location of the mobile part, and the location identifying means of the mobile part is arranged to respond to a location request from the interrogation means with a non-unique location signal which, in combination with the approximate location determined by the fixed part, determines a unique location.

In a preferred arrangement, the fixed part and the mobile parts each have a satellite navigation system receiver, and the positions of the mobile parts as measured by the satellite navigation system are compared with those of the fixed part as measured by the satellite navigation system. The position of the fixed part can be known with great accuracy and provides a reference measurement which allows the position of the mobile part to be determined with greater accuracy than is possible by direct measurement using the satellite system alone.

Preferably the fixed part has one or more servers and means for allocating a server to a mobile part only when it requires service. In practice only a very small number of mobile units will require service at any given time, so this allows the computing resources of the fixed part to be used most efficiently, and the system can support many more mobile units in total than it has server capacity for. This is in contrast to the prior art system discussed above, in which each mobile unit requires a dedicated computer carried on board, which is only used for a fraction of the time. Moreover, all the servers can use a common road-use database, which can use the information on routes it has planned for mobile users to build a prediction of future road use status, such as likely congestion points, and build this into its guidance instruction process. For example the system can be arranged such that it does not direct more than a predetermined number of users to use a particular stretch of road at a particular time, and finds alternative routes for any users who would otherwise be directed along that road at that time. In this way the system can predict likely congestion points and take pre-emptive action.

The mobile part may include guidance instruction means controllable by instructions contained in the guidance information transmitted from the fixed part over the communications link, whereby guidance instructions can be communicated to the user by means of the guidance instruction means.

For some applications the vehicle may be controlled directly in response to the guidance information received over the communications link. However, for use on the public highway, it is preferable that the guidance information controls display means, which may be visual or audible or both, to indicate to a driver the direction to take.

The guidance instruction means may be programmable from the fixed part over the communications link, either automatically or by a human operator. The guidance instruction means may include a speech synthesiser, which may be located in the fixed part, transmitting voice messages to the user over the communications system, or may be located in the mobile unit and controlled by data messages from the fixed part. The former arrangement allows the mobile unit to be simplified, whilst the latter arrangement requires a smaller signalling load.

In the described embodiment the mobile part is in a vehicle, but it may be a hand-held device for guiding a pedestrian. In one form, the mobile part may be a conventional mobile cellular radio unit. This allows a basic service to be provide to a user without the need for any dedicated equipment.

Embodiments of the invention will now be described by way of example with reference to the drawings, in which:

Figure 1 shows a mobile part and a fixed part of a navigation information system according to an embodiment of the invention;

Figure 2 illustrates how the invention may be applied to a simple road layout;

Figure 3 illustrates the division of a territory into zones according to the instructions generated by the system;

Figure 4 illustrates an application of the invention to a more complex road layout;

Figures 5a and 5b illustrate the modification of an overlay in response to a change in traffic circumstances; and

Figure 6 illustrates a road network, showing overlay areas defined by the method of the invention in relation to a cellular radio network

According to the embodiment of Figure 1 the navigation system has a fixed part (comprising elements 12 to 19) and a number of mobile parts, of which one only is shown (comprising elements 1 to 10), interconnected by a cellular telephone network 11.

The mobile part comprises a mobile telephone 1 having an audio output 2, an audio input 3 and a radio antenna (transmit/receive) 4. The output 2 is connected to a decoder 5 to translate Dual-Tone Multi-Frequency (DTMF) signals received by the telephone 1 into data which is fed to an interface controller 6. The interface controller 6 also receives input from a GPS (Global Positioning System) satellite receiver 7. The interface controller transmits data to a DTMF encoder 8 which generates tones to be fed to the audio input of the mobile telephone. The audio output 2 and input 3 also include a loudspeaker 9 and microphone 10 respectively, to allow the telephone to be used for speech.

The fixed part comprises an interface with the cellular telephone network 11, connected through a DTMF decoder 12 and encoder 13 and a controller interface 14 to a computer 15. The computer 15 comprises a number of servers 16, one of which is allocated to each active mobile unit. The servers 16 have access to a geographical database 17, and a database of standard messages 18. The geographical database 17 is updateable through updating input 19. The database 17 stores the definitions of a number of overlay areas which together form a geographical overlay to the territory to be covered. Examples of overlays are illustrated in Figures 2, 4, 5a, 5b, and 6, to be described in detail later.

The mobile part obtains location information using the GPS receiver 7 and transmits this information, together with a request for directions to a specified destination, to the fixed part, where a server 16 relates the location information to its geographical database 17 and obtains message information associated with the location from the database 18, and transmits the information back to the mobile part.

The computer 15 may transmit messages in DTMF code, using the encoder 12, or it may generate voice messages which are transmitted through a voice output 20 to the cellular network 11.

DTM= signals are used to transmit the position of the vehicle to the computer 15 which can then offer information and guidance either to the vehicle or to a third party on demand.

In the following discussion, variations on the basic apparatus depicted in Figure 1 will also be described, in which certain elements are modified or replaced.

The system is operated as follows:-

At the start of a journey the driver requests service by activating a pre-dialled control on the telephone 1. This service request is transmitted to the control interface 14 over the telephone network 11. The control interface 14 then allocates a free server 16 to answer the call and interrogate the vehicle GPS receiver 7 to determine its geographical position. The encoder 8 takes the latitude and longitude data and translates the numbers into DTMF tone-pairs, in a manner to be described in more detail below.

The cellular telephone couples this audio signal into its speech input path. This is easy to do with a hands-free vehicle-mounted cellular telephone since the microphone lead is accessible or alternatively, a small transducer can be mounted next to the microphone 10. A DTMF receiver 5 coupled to the loudspeaker 9 (again acoustically or electrically) decodes supervisory data (again in DTMF format) coming back from the server 16 to acknowledge the reception of location messages. If no acknowledgement is received by the DTMF unit then the data message is repeated.

The fixed end of the system comprises a DTMF decoder 12 and encoder 13 coupled to a serial data interface 14 of the server computer 15. This computer, on the one hand, can call the mobile part which will answer automatically and then provide its location using the DTMF signalling system or on the other hand can receive an unsolicited call, which would include the DTMF encoded identity of the mobile unit and would also provide the vehicle location using the DTMF interface 6.

The server 16 then captures the current position of the user, and identifies the overlay area within which that position falls. The server also captures any permanent user-specific information such as the type of vehicle, which may be relevant for the route to be selected e.g. because of height or weight restrictions. The user may encode those requirements which are not permanent, but are

specific to the present information request, (in particular his destination) by using the telephone keypad in response to voice prompts. However, in a preferred arrangement the call is presented to a human operator for the capture of this data. This allows the user to obtain assistance in identifying his desired destination to the system, and also allows the driver to speak his requirements, keeping his hands and eyes free for driving.

The operator then remotely programs the in-vehicle interface 6 with system data identifying the vehicle destination, for use in subsequent update processes, and instigates the generation of voice given directions and instructions to the driver by a speech generation subsystem of the computer server 16.

Position fixes may be made at regular intervals, e.g. every two minutes, or every kilometre. Alternatively the fixed part may request the mobile unit to send its next position fix after a specified interval or distance.

As the driver follows the route further instructions can automatically be sent as the driver enters each new overlay area and the driver can be alerted if the route has been left or if any new traffic problems have been detected that will affect the individual driver. The system is arranged such that when the system locates a mobile unit entering an overlay area having a message defined for it, for example the next turn instruction (or an error message if the mobile unit has gone off the selected route), that message is transmitted. The system may also be arranged to transmit messages to users other than the mobile unit in question, for example to monitor the progress of valuable cargoes.

At any time the driver can call the human operator if service requirements change or additional help is needed.

Because a central database is used all vehicle movements can be monitored. Traffic models can be used to optimise traffic flows and reduce journey times. The system can also ensure that it does not itself cause congestion, by limiting the number of vehicles it directs to use the same road at the same time. The control system can use the location data to calculate and record movement vectors from these vehicles.

Using the data collected by this method, it is possible for the central system to derive a digital map of valid routes. The following data could be derived automatically: valid travel lanes; permitted direction(s) of flow; allowable turns;

average travel times; trends in travel times according to time of day and other factors.

The system would automatically update the map to show permanent changes (new road links, changes to one way systems etc.). Temporary lane closures from road works etc. would also be recorded. Manual updating of data would be necessary (for instance to alert the system to a new bypass opening) before the system acquired the information from vehicle flow data, to ensure vehicles are routed over the new road initially. Any approximations in the pre-entered data would automatically be corrected by the system described here.

The system could be further enhanced to include any other information that may be relevant to travellers, by a combination of manual and automated data entry, e.g. location of bus stops, telephone boxes and other street furniture, and proximity to enterprises such as shops, banks or offices.

The variation of transit time trends according to time of day, for each link, could be used to derive a congestion prediction model, as the basis for route guidance. The system may monitor the progress of the mobile units along the routes selected for them, to identify any areas of traffic congestion etc, by comparing actual transit times between predetermined locations. This may be done by the fixed system monitoring the location updates of individual units, or it may be done by the mobile unit, in co-operation with the fixed unit. In this latter case, the fixed part transmits an expected range of transit times within which the mobile is expected to reach a predetermined location. If the mobile unit reaches the location outside this range, it reports the fact to the fixed part. By "reporting by exception" the data processing overhead can be reduced considerably.

However, these systems can become unstable if too many drivers have access to route guidance based on information about current or predicted congestion. To avoid these instabilities route plans are created and updated centrally and passed to individual vehicles. The impact of these vehicles using the suggested routes is then added to the prediction. As more vehicles use the system the prediction produced could become more accurate.

The routes derived can be passed to the vehicles (via a mobile data link, or possibly a short range communications link or other temporary access to a fixed

telecommunications network - prior to departure). The vehicle would then operate autonomously, unless the road conditions varied significantly from those predicted.

If the central system detected a problem (from vehicle data or other sources), which had a severe impact on predictions, sufficient to cause a change to advice already given, then the central system could broadcast news of the problem, such that those vehicles affected could automatically call in via a mobile data communications link to receive a new route from its present location to its destination.

If a vehicle system encountered unexpected transit times along its programmed route it would send a report to the central system.

The data flowing though the system will therefore allow it to "learn" more of the road network's characteristic congestion behaviour, e.g. by use of neural net techniques, and to select routes for traffic which avoid using routes at times when they are likely to be congested. In addition, the system can generate digital road maps or other data automatically, based on the position measurements of vehicles using the roads.

A particular advantage of this system is the ability to predict unusual patterns of congestion from the route guidance information requested by the users. Because route guidance is generated centrally, the system can monitor the number of requests for destination information to a given location. By determining the predicted arrival times for each user (which will depend on their starting points, and the time the journey started), a build-up of traffic converging on a particular location at a particular future time (e.g. for a major sporting event) can be detected. Traffic for other destinations, which might have been routed by way of this location, can then be diverted to other routes.

The system described above uses an analogue telecommunications link, in which DTMF codes may be used. For an analogue cellular radio network DTMF is an ideal signalling medium when only short status messages are required to be transmitted. It can survive in the severe signal fading and noise of the mobile environment which frequently precludes the use of fast phase or frequency shift data modulation. Another advantage is the ability to co-exist with speech. For example a DTMF data burst containing vehicle position data could be sent at the start of a call and at intervals during the call. Other simple coded DTMF

messages can also be conveyed to indicate emergencies, provide simple driver indications (e.g. illuminated arrows to turn left or right) or trigger synthetic speech generated by another sub-system in the vehicle.

The DTMF coding described above is suitable for an analogue system. In a digital cellular network digitised data can be transmitted over an associated packet data system such as the Short Message Service (SMS) of GSM (Global System for Mobile Communications), or the General Packet Radio Service (GPRS) proposed for GSM.

In the embodiment described above, the speech generation subsystem forms part of the server 16. Alternatively, it can be carried on board the vehicle. In this arrangement the subsystem has various stored speech commands which are controlled from the in-vehicle interface 6 in response to commands transmitted from the fixed part. This arrangement reduces the signalling traffic required over the radio link 11, but increases the complexity of the in-vehicle equipment.

The location-determination system will now be described in greater detail. GPS (Global Positioning System) satellite navigation receivers are now becoming very cheap and are available with a serial data output. These can provide latitude and longitude data to within a tenth of a second of arc (defining position to within 3 metres, which is sufficient to identify which carriageway of a dual carriageway road a user is on).

Satellite positioning systems such as the Global Positioning System (GPS) are prone to small systematic errors, for example as a result of instabilities in the orbits of the satellites. The accuracy of the position measurement may be enhanced by a process known as "Differential GPS" in which a number of fixed reference points are used, whose positions are determined with great precision e.g. using surveying techniques. GPS is used to obtain a measure of the position of one or more of the fixed reference points. This measure is compared with the known, true location to generate a correction value which can be used to correct the position of the mobile unit as measured by GPS.

The position data received from the satellite positioning system may include some redundant data. If the system is only to operate within a limited area of the globe the most significant digits of the position data are redundant, and need not be transmitted from the mobile unit to the fixed part. For example, any

point in Germany can be uniquely defined by the units digits of its latitude and of its longitude, as that country lies entirely between 45 and 55 degrees North, and between 5 and 15 degrees East. It is also possible to define any point in the United Kingdom in this way, although in that case a 10 degree offset in longitude has to be applied to avoid duplication of longitudes East and West of the zero meridian.

For larger territories e.g. a pan-European system, or one covering the USA, this simple method of data reduction is impractical. However, it is nevertheless possible to reduce the data requirements by dynamically defining the territory. After an initialisation step using the full location, the system selects as each new location the closest candidate to the previous one. For example, if the mobile unit was last reported at 99 degrees W and the units digit of the longitude is now 0, the user is taken to be at 100 degrees W rather than, for example, 90 degrees or 110 degrees.

If location updates take place sufficiently frequently that the user's position cannot have changed by more than half a degree, the units digit of degrees may also be dispensed with, and the location given only in minutes and seconds of arc. The more frequent the updates, the more digits can be dispensed with.

An alternative method of obtaining the coarse position location is interrogation of the cellular radio system's operating system to identify the cell in which the user is currently located. Cell sizes can be up to about 40km across (although they are often much smaller, so identifying the cell can identify the user's location to within 40km, which identifies latitude to better than half a degree. (1 degree of latitude = 111km). The separation of lines of longitude varies with the cosine of the latitude but even at the Arctic Circle (66 degrees North) a 40km resolution will identify longitude to the nearest whole degree (1 degree of longitude = 111km (cos latitude) = approximately 45km at 66 degrees North).

By left-truncating the position data by omitting the degrees digits a basic position message would therefore consist of 10 decimal digits (minutes, seconds, and tenths of seconds). Altitude data giving altitude in metres would require a further four digits, since all points on the Earth's surface lie within a range of 10,000 metres, but this data can also be left-truncated, as it is unlikely that any multi-level road system would exceed 100 metres in height (or if it did, that a GPS

system would work effectively for any receiver on the lower levels). This gives a total of twelve digits, which can be transmitted by DTMF in less than 2 seconds.

If the data is left-truncated as described above, the "coarse" data is added by the interface controller 14 by reference to the previous position or to the cellular radio operating system.

When the computer 15 receives a location message, it stores the location and then searches its database for an overlay area within which that position lies. The overlay areas are defined in the database by co-ordinates of latitude and longitude and have associated attributes which define messages which can be passed to mobile subscribers within the overlay area defined. In some instances height (altitude) information, also available using satellite positioning systems, may be used, for example to distinguish between levels in a multi-level highway intersection. When a DTMF location message has co-ordinates which fall inside an overlay area having an associated message, the message is then transmitted to the mobile part as a computer synthesised speech message, a DTMF coded message (to activate other subsystems) or as a high speed conventional data message.

If the mobile unit fell within the same overlay area at the previous location update, and the message associated with that overlay area is unchanged, the transmission of the message may be suspended.

The frequency at which location updates are requested by the system may be tailored to the size and nature of the current overlay area. For example, an intricate road layout may comprise a large number of small overlay areas, requiring frequent location updates to ensure that a user does not miss an instruction by passing through its associated area between two updates. However, a long stretch of road without junctions may be covered by a single overlay area, so less frequent updates are appropriate. The speed with which a vehicle is likely to be moving, which will differ between urban, rural, and motorway environments may also be used as a factor in determining when the next location update should be requested.

As suggested above, there may be circumstances when a satellite positioning system may be unusable, for example in tunnels or built-up areas where a line-of-sight view of the satellites may be impossible to obtain. Alternative arrangements for identifying and updating the mobile part's location which do not

rely on a satellite receiver may be used, either on their own, or to interpolate between points where a satellite system can be used. In one variant, a navigation system based on dead-reckoning may be used. In such systems the user identifies his initial location and the on-board system measures the system's movement e.g. by magnetic bearing measurements, distance counters, and inertial navigation means such as gyrocompasses and accelerometers. Such systems are self-contained, but require knowledge of the starting point. This may be obtained, for example from a satellite positioning system.

In another variant, a method of location may be used which relies on the propagation characteristics of the cellular radio system used for communication with the central control station. Examples of such systems are disclosed in German Patent specifications DE3825661 (Licentia Patent Verwaltungs) and DE 3516357 (Bosch), United States Patent 4210913 (Newhouse), European Patent specification EP0320913 (Nokia), and International Patent applications WO92/13284 (Song) and WO 88/01061 (Ventana). By comparison of signal strength or other characteristics of several cellular base stations, a position fix can be determined. In this arrangement the location measurement may be made directly by the fixed system. This allows the mobile part of the system to be embodied by a conventional cellular telephone, with inputs being provided by speech, or by DTMF tones generated by the keypad, and instructions to the user being transmitted by voice commands.

Examples of the kind of navigation information which may be stored in the database 17 will now be discussed, with reference to Figures 2 to 6. Briefly, Figure 2 shows a junction J having four approach roads 21,22,23, 24; each having associated with it an overlay area 21a, 22a, 23a, 24a respectively. In this figure, and all other figures illustrating road layouts, the roads are shown arranged for left-hand running, as used for example in the UK, Japan, Australia etc. Figure 3 shows part of a road network surrounding the junction J, including towns A, B, C, and a motorway M. Each of the roads 21, 22, 23, 24 has an associated destination zone 21z etc. Figure 4 shows a complex grade-separated junction interlinking four roads N, S, E, W. The junction has superimposed on it an overlay having twelve overlay areas, Na, Ni, Nd, Sa, Si, Sd, Ea, Ei, Ed, Wa, Wi, Wd. Figure 5a shows a small region having a main road 33 and a side road 30. The main road 33 has two

associated overlay areas 31, 32. Figure 5b is similar to Figure 5a, but an obstruction X is present on the main road 33, and the overlay area 32 has been subdivided into two overlay areas 32a, 32b, separated by the obstruction. Figure 6 shows an overlay comprising ten overlay areas 40 - 49 superimposed on a cellular radio coverage region comprising five cells 50 - 54.

In greater detail, the road junction J (Figure 2) has four approach roads 21, 22, 23, 24. On each road, at the approach to the junction, an overlay area (21a, 22a, 23a, 24a) is defined. These overlay areas have directional information associated with them, giving turn instructions or other navigational information. As shown in Figure 3, the entire territory covered by the navigation system can be divided into four zones 21z, 22z, 23z, 24z, each comprising the set of all locations for which the corresponding road 21, 22, 23, 24 should be taken from the junction J. In this particular example, road 24 leads directly into town A and is only used for local destinations (zone 24z), road 23 leads to town B (zone 23z), road 22 leads to town D (zone 22z) and road 21 leads to the motorway M, for all other destinations including town C and part of town A. These zones are defined differently for each junction: for example at junction J' different directions are appropriate for towns A and C, so these towns fall in different zones with respect to the overlay areas at that junction. The zones may even be defined differently for different overlay areas at the same junction. For example, if U-turns are not possible at the junction J, any traffic approaching the junction J by road 22 and requiring town D (perhaps as the result of a previous error, or a change of plan) must be routed by way of roads 21, M, and 25. Thus, for overlay area 22a there are only three zones: 24z, 23z and the combined 21z/22z, corresponding to the three permitted exits 21, 23, 24.

The zones may be re-defined according to circumstances. For example, when the motorway M is congested, the best route from junction J to town C may be by way of town B. In such circumstances, zones 21z and 23z are redefined so that town C now falls within zone 23z. It should be noted, however, that the total number of zones remains the number of exit routes from the relevant overlay area.

The overlay areas 21a, 22a, 23a, and 24a should be large enough to ensure that any vehicle approaching the junction gets at least one location update whilst within the relevant overlay area, and is thus sent the relevant turn

instruction. As shown in Figure 2, these overlay areas are discrete, and may be considered equivalent to the coverage areas of the beacons of the prior art system discussed above. They may, however, be made contiguous, as shown in Figures 4, 5a, 5b and 6.

Figure 4 shows a more complex, grade-separated junction, in which there are twelve overlay areas. Each road N, E, S, W intersecting at the junction has a corresponding approach overlay area Na, Ea, Sa, Wa, (Wa shown shaded), and a depart overlay area Nd, Ed, Sd, Wd (Ed shown shaded). There are also four intermediate overlay areas Ni, Ei, Si, Wi (Si shown shaded). In the vicinity of the flyover F height (altitude) information obtainable from the GPS system can be used to determine which level, and therefore which overlay area, the user is currently in.

The approach and intermediate overlay areas each end at a decision point P1 to P8. In the database 17 each overlay area has direction information associated with it, providing instructions as to which fork to take at the associated decision point. For example, the direction information associated with zone Si instructs users for destinations served by road N to go straight on at point P1, and users for destinations served by roads E, S, and W to turn left. It will be seen that traffic using the intersection will pass through one approach overlay area, one departure overlay area, and may also pass through one or more intermediate overlay areas. There may also be information associated with the departure overlay areas Nd, Sd, Ed, Wd, for example warning of hazards ahead. The departure overlay areas may be continuous with approach overlay areas for the next junction in each direction.

As a user approaches the junction on road S, a location update identifies the user equipment as being within overlay area Sa. If the co-ordinates of the user's destination are within the zone served by road W, the user is sent an instruction to turn left at point P2. If the user obeys this instruction, he will enter overlay area Wd and on the next location update he will be sent information relevant to that overlay area (if any).

If the co-ordinates of the user's destination are within the zone served by road N, the user in overlay area Sa is instead sent an instruction to continue straight on at point P2. If the user obeys this instruction, he will enter overlay area Si.

For a user in overlay area Si, if the co-ordinates of the user's destination are within the zone served by road N the user is sent an instruction to go straight on at point P1. On obeying this instruction, he will enter the overlay area Nd and on the next location update he will be sent information relevant to that overlay area (if any).

If the co-ordinates of the destination of a user in overlay area Si are in the zone served by roads E, S, or W, the user will be sent an instruction to turn left at point P1. On obeying this instruction, he will enter overlay area Wi.

Similar information is associated with the other overlay areas. By being given appropriate instructions as the user negotiates a succession of junctions (decision points), the user can be directed to any destination. It should be noted that all users who are to be directed to the same exit from the junction are given the same instruction, whatever their ultimate destination.

Figures 5a and 5b illustrate the reconfiguration of the overlay areas to meet changing circumstances. Initially (Figure 5a) an overlay area 31 is defined for the approach to a junction between a major road 33 and a side road 30, and a second overlay area 32 is defined for that part of the major road 33 beyond the junction. Information associated with the overlay area 31 includes turn information to instruct traffic for the zone served by the side road 30 to turn off. Information may also be associated with the overlay area 32.

In figure 5b the major road 33 has been blocked at a point X. In order to accommodate this, the overlay area 32 has been subdivided into two overlay areas 32a, 32b. The information (if any) associated with overlay area 32b is the same as that previously associated with overlay area 32. Traffic in overlay area 32a is given new information warning it of the hazard ahead. The information associated with the overlay area 31 is modified, so that all traffic is now instructed to turn off onto the side road 30. (Effectively this means that the destination zones associated with the overlay area 31 are merged into one)

Figure 6 shows how the overlay areas may be defined for a road network. In this example there is an overlay area 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 corresponding to each side of each section of road. Information appropriate to each direction of travel on each section is therefore available to users throughout the relevant section. Superimposed on this overlay there is a cellular radio network,

five cells of which (50, 51, 52, 53, 54) are shown. The position of the user, as determined for example by a satellite positioning system, determines which overlay area is appropriate to the user. The information is transmitted to the service control centre by means of the cellular radio network. Handovers between cellular base stations occur in conventional manner at cell boundaries. These handovers are, however, unrelated to the boundaries between the overlay areas 40 - 49

Although the described embodiment relates to the provision of route guidance information, other locality-dependant information may be provided as well, or instead, such as information about local facilities, tourist attractions, weather forecasts, public transport information, etc. The term "guidance information", as used in this specification, embraces any such information.

CLAIMS

1. A navigation information system for providing information to a mobile user dependant on the location of the mobile user, the system comprising a mobile communications system having a fixed part and one or more mobile part for communicating with the fixed part, the one or more mobile part including means for transmitting to the fixed part a request for guidance information and for receiving guidance information from the fixed part, and the fixed part including:

means for determining the location of a mobile part requesting guidance information,

means for generating guidance information according to the location of the mobile part, and

means for transmitting the guidance information so generated to the mobile part,

whereby information dependant on the location of the mobile unit can be transmitted to the mobile unit.

2. A system as claimed in Claim 1, the fixed part including means for determining the location of the mobile part in relation to a geographical overlay comprising a plurality of overlay areas, and means for transmitting information associated with an overlay area which includes the location of the mobile part, whereby a mobile part within that overlay area receives information associated with that overlay area.

3. A system as claimed in Claim 2, including means for storing a digital representation of the geographical overlay, and means for modifying the stored representation such that the configurations of the overlay areas may be selected to meet changing requirements.

4. A system according to Claim 2 or 3, including means for determining when a mobile part enters a predetermined overlay area, and means for transmitting a message to the mobile part in response to the mobile part entering the predetermined overlay area.

5. A system according to Claim 2, 3, or 4 including means for determining when a mobile part enters a predetermined overlay area, and means for transmitting a message, to a user other than the said mobile part, in response to the said mobile part entering the predetermined overlay area.

6. A system according to claim 4 or 5, including means to store a value associated with the mobile part, and means arranged to modify the stored value in response to the message.

7. A system as claimed in any preceding claim, having means for locating the position of the mobile part by radio location

8 A system as claimed in Claim 7, wherein the means for locating position comprises a satellite navigation system receiver and/or means for identifying the location of the mobile part in relation to elements of the fixed part of the communications system.

9. A system as claimed in any preceding claim, wherein the means for determining the location of the mobile part comprises means to interrogate a location-identifying means forming part of the mobile part.

10. A system as claimed in claim 9, wherein the fixed part has means to determine the approximate location of the mobile part, and wherein the location identifying means of the mobile part is arranged to respond to a location request from the interrogation means with a non-unique location signal which, in combination with the approximate location determined by the fixed part, determines a unique location.

11. A system as claimed in any preceding claim, wherein the mobile part has means for locating its position by dead reckoning

12. A system as claimed in any preceding claim, the fixed part including means for generating and maintaining guidance data based on vehicle movement data derived from time information and position measurements of a plurality of the mobile parts and/or estimations of future locations of the mobile parts based on the guidance information previously transmitted to the mobile parts.

13 A system according to any preceding claim wherein the fixed part comprises means for transmitting to the mobile part an expected range of movement information and for receiving from the mobile part movement measurements outside the expected range, and the mobile part comprises means for measuring location and time to derive movement information, means to compare the movement information with the expected range received from a fixed part of the system, and means to automatically report to the fixed system movement measurements outside the expected range.

14. A system according to any preceding claim, the fixed part including means for storing guidance data, means for updating the stored guidance data, means for identifying mobile parts to which the updated data are applicable, and means for transmitting such data over the communications system to the mobile parts so identified.

15. A system according to any preceding claim, wherein the mobile part includes guidance instruction means controllable by instructions contained in the guidance information transmitted from the fixed part over the communications link, whereby guidance instructions can be communicated to the user by means of the guidance instruction means.

16. A system according to any preceding claim, wherein the fixed part has input means operable by a human operator to input guidance instruction requests to the fixed part.

17. A navigation information system for providing information to one or more mobile user dependant on the location of the one or more mobile user, the system comprising:

means for determining the location of a mobile unit requesting guidance information,

means for generating guidance information according to the location of the mobile unit,

and a communications system for transmitting the guidance information so generated to the mobile unit,

whereby information dependant on the location of the mobile unit can be transmitted to the mobile unit.

18. A system as claimed in Claim 17, including means for determining the location of a mobile unit in relation to a geographical overlay comprising a plurality of overlay areas, and means for transmitting information associated with an overlay area which includes the location of the mobile unit, whereby a mobile part within that overlay area receives information associated with that overlay area.

19. A system as claimed in Claim 18, including means for storing a digital representation of the geographical overlay, and means for modifying the stored representation such that the configurations of the overlay areas may be selected to meet changing requirements.

20. A system according to Claim 18 or 19, including means for determining when a mobile unit enters a predetermined overlay area, and means for transmitting a message to the mobile unit in response to the mobile unit entering the predetermined overlay area.

21. A system according to Claim 18, 19, or 20 including means for determining when a mobile unit enters a predetermined overlay area, and means for transmitting a message, to a user other than the said mobile unit, in response to the said mobile unit entering the predetermined overlay area.

22. A system according to claim 20 or 21, including means to store a value associated with the mobile unit, and means arranged to modify the stored value in response to the message.

23. A system as claimed in any of claims 17 to 22, the means for determining the location of a mobile unit comprising means to interrogate a location-identifying means of a co-operating mobile unit to determine its position..

24. A system as claimed in any of claims 17 to 23, wherein the means for locating position comprises means for identifying the location of the mobile unit in relation to elements of the fixed part of the communications system.

25. A system as claimed in claim 24, wherein the means for locating position comprises means to determine the approximate location of the mobile unit, means to receive a non-unique location signal from the mobile unit, and means to combine the approximate location information with the non-unique location information to determine a unique location.

26. A system as claimed in any of claims 17 to 25, including means for generating and maintaining guidance data based on vehicle movement data derived from time information and position measurements of a plurality of the mobile parts

27. A system according to any of claims 17 to 26, having means for transmitting to the mobile part an expected range of movement information, and for receiving from the mobile part movement measurements outside the expected range.

28. A system according to any of claims 17 to 27, including means for storing guidance data, means for updating the stored guidance data, means for identifying mobile units to which the updated data are applicable, and means for transmitting such data over the communications system to the mobile units so identified.

29. A system as claimed in any of claims 17 to 28, having input means operable by a human operator to input guidance instruction requests.

30. A mobile unit for a navigation information system, comprising means for receiving guidance instruction information over a communications link, and guidance instruction means controllable by the guidance instruction information received over the communications link, whereby guidance instructions can be communicated to the user by means of the guidance instruction means.

31 A mobile unit for a navigation information system comprising means for measuring location and time to derive movement information, means to compare the movement information with an expected range received from a fixed part of the system, and means to automatically report to the fixed system movement measurements outside the expected range.

32 A method of providing navigation information to mobile units of a mobile radio system dependant on the locations of the mobile units comprising the steps of storing navigation data in a fixed part, transmitting a request for navigation guidance from a mobile unit to the fixed part, determining the location of the mobile unit, generating guidance information on the basis of the stored data, location information and the request, and transmitting the guidance information from the fixed part to the mobile unit, whereby information relevant to the location of the mobile unit is transmitted to the mobile unit.

33. A method as claimed in Claim 32, wherein the location of the mobile unit is determined in relation to a geographical overlay comprising a plurality of overlay areas, generating information associated with an overlay area which includes the location of the mobile part, and transmitting the information associated with the

relevant overlay area to the mobile part, whereby a mobile part within that overlay area receives information associated with that overlay area.

34. A method as claimed in claim 33, including the step of storing a digital representation of the geographical overlay, and modifying the stored representation such that the configurations of the overlay areas may be selected to meet changing requirements.

35. A method according to Claim 33 or 34, comprising the further steps of determining when a mobile unit enters a predetermined overlay area, and transmitting a message to the mobile unit in response to the mobile unit entering the predetermined overlay area.

36. A method according to Claim 33, 34, or 35 including the further steps of determining when a mobile unit enters a predetermined overlay area, and transmitting a message to a user other than the said mobile unit in response to the mobile unit entering the predetermined overlay area.

37. A method according to claim 35 or 36 including the further step of modifying a stored value associated with the mobile unit in response to the message.

38. A method as claimed in any of claims 32 to 37, wherein the position of the mobile unit is identified by a radio location method.

39. A method as claimed in Claim 38, wherein the position of the mobile unit is determined by means of a satellite navigation system and/or by identifying the location of the mobile part in relation to elements of the fixed part of the communications system.

40. A method according to any of Claims 32 to 39, wherein the fixed unit interrogates the mobile unit to identify its location.

41. A method as claimed in claim 40 wherein the fixed part determines the approximate location of the mobile part, and wherein the mobile part responds to a location request from the interrogation means with a non-unique location signal which, in combination with the approximate location determined by the fixed part, determines a unique location.

42. A method as claimed in any of claims 32 to 41, wherein the mobile unit identifies its position by dead reckoning.

43. A method according to any of Claims 32 to 42, including the steps of generating and maintaining data based on vehicle movement data derived from time information and position measurements of a plurality of the mobile parts and/or estimations of future locations of the mobile parts based on the guidance information previously transmitted to the mobile parts.

44. A method according to any of claims 32 to 43 wherein the fixed part transmits to the mobile part an expected range of movement information, and the mobile part measures location and time to derive movement information, compares the movement information with the expected range received from the fixed part of the system, and reports to the fixed system movement measurements outside the expected range.

45. A method as claimed in any of Claims 32 to 44 including the further steps of the updating the stored data, identifying the mobile units to which the updated data are applicable, and transmitting such data over the communications system to said applicable mobile parts.

46. A method as claimed in any of claims 32 to 45, wherein the guidance information transmitted to the mobile unit controls guidance instruction means forming part of the mobile unit, whereby guidance instructions can be communicated to the user of the mobile unit.

47. Apparatus substantially as described with reference to the accompanying drawings

48. A method substantially as described with reference to the accompanying drawings

Fig.1.

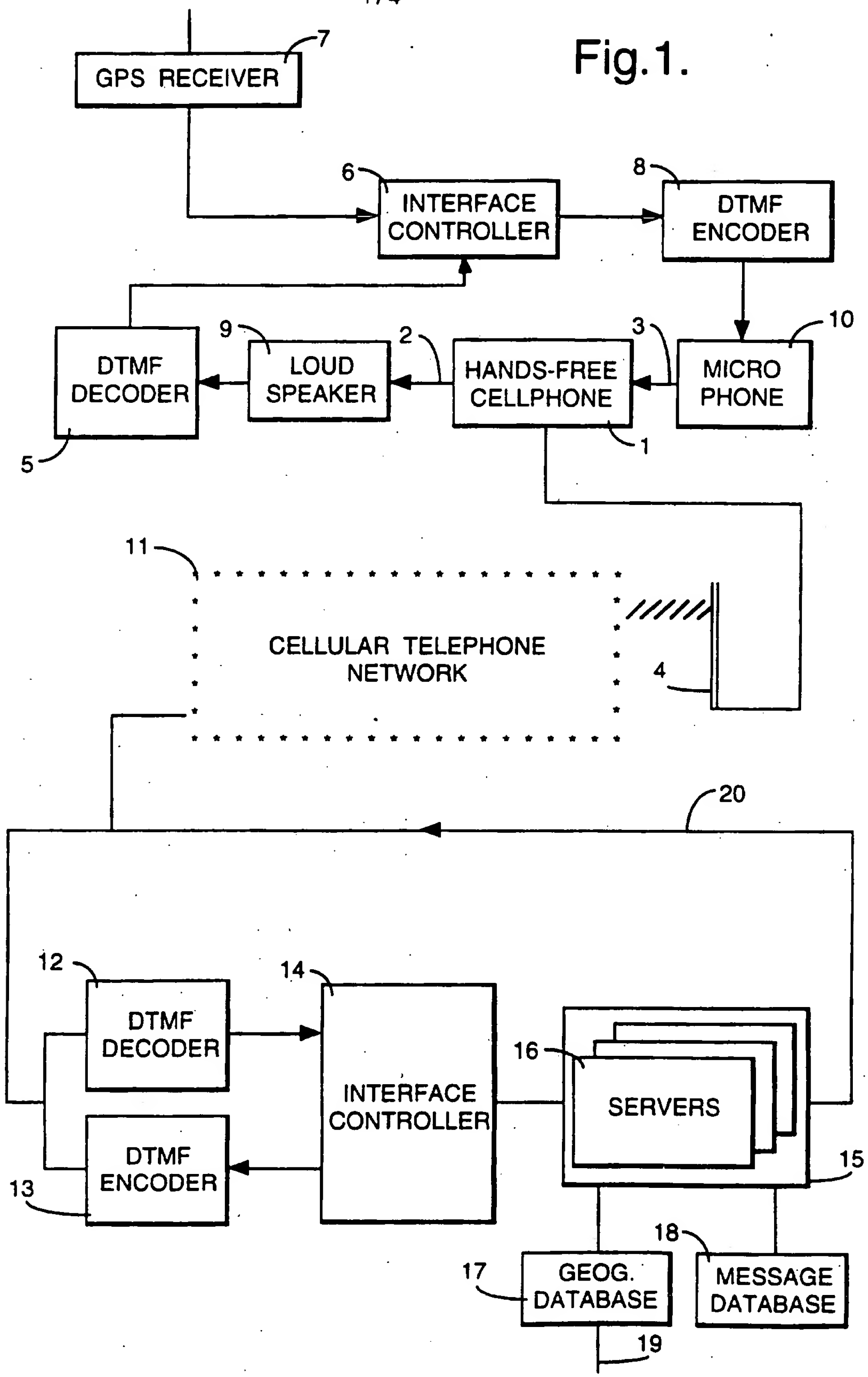

GPS RECEIVER
7
6
INTERFACE CONTROLLER
8
DTMF ENCODER
9
LOUD SPEAKER
2
3
10
DTMF DECODER
HANDS-FREE CELLPHONE
MICRO PHONE
1
5
11
CELLULAR TELEPHONE NETWORK
4
20
12
14
DTMF DECODER
16
INTERFACE CONTROLLER
SERVERS
DTMF ENCODER
15
13
18
17
GEOG. DATABASE
MESSAGE DATABASE
19

Fig.2.

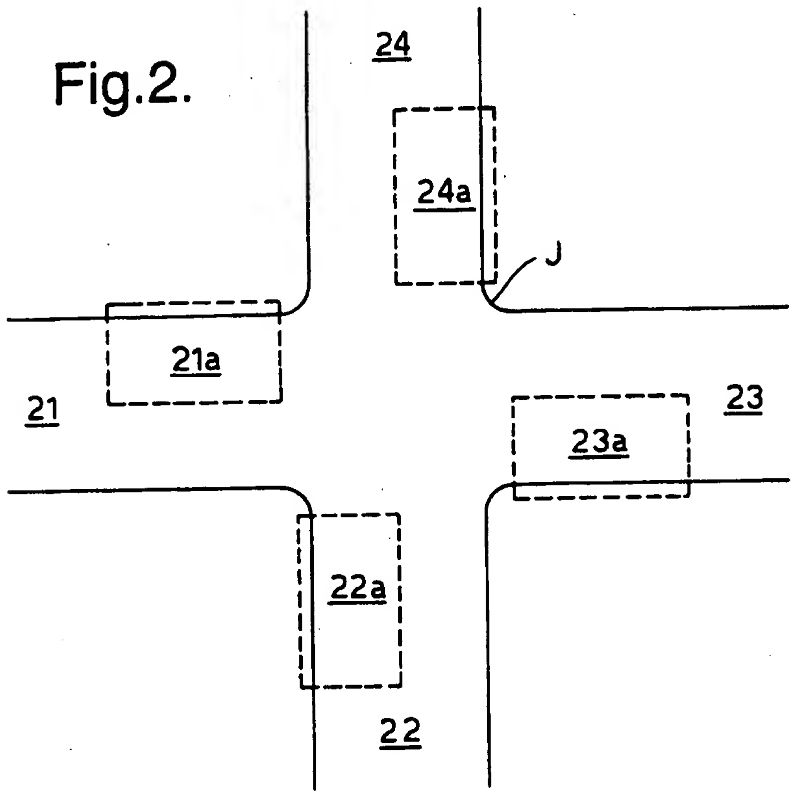

24
24a
J
21a
21
23
23a
22a
22

Fig.3.

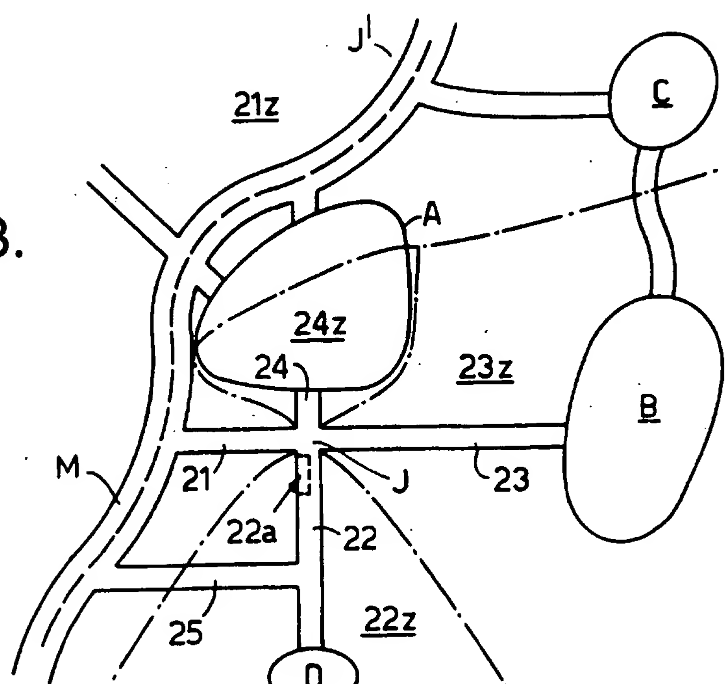

J'
21z
C
A
24z
24
23z
B
M
21
J
23
22a
22
25
22z
D

Fig.4.

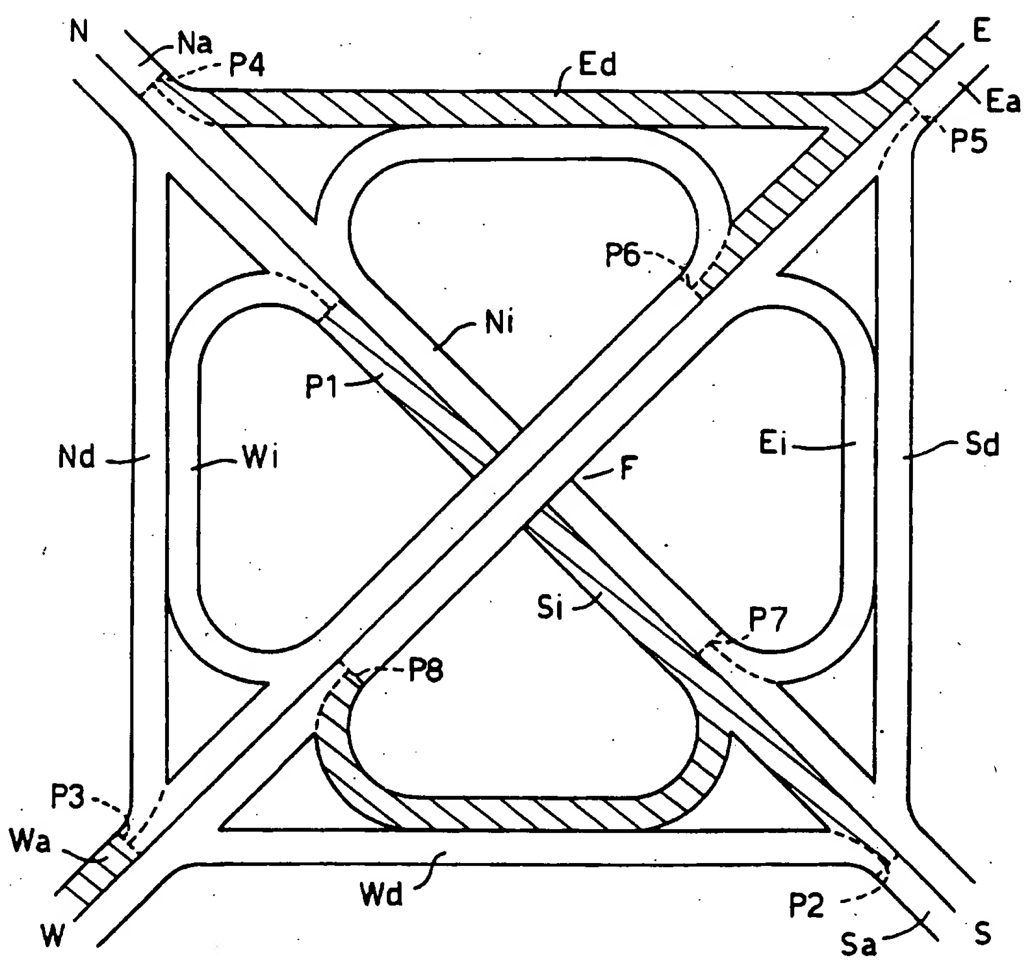


Fig.5a.

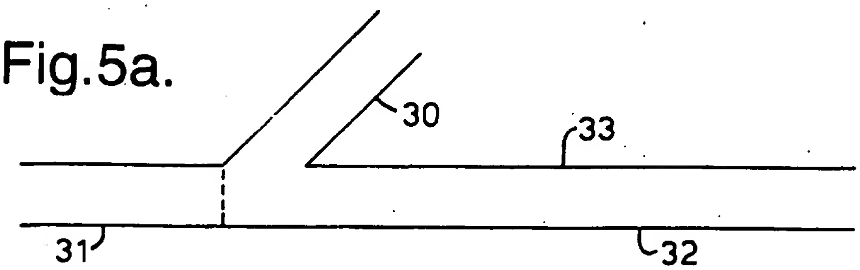


Fig.5b.

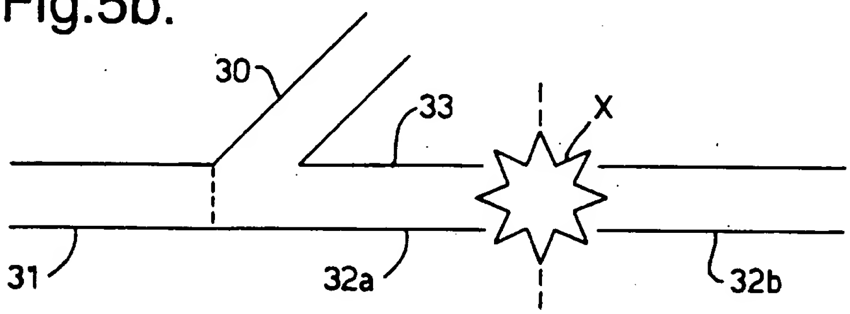
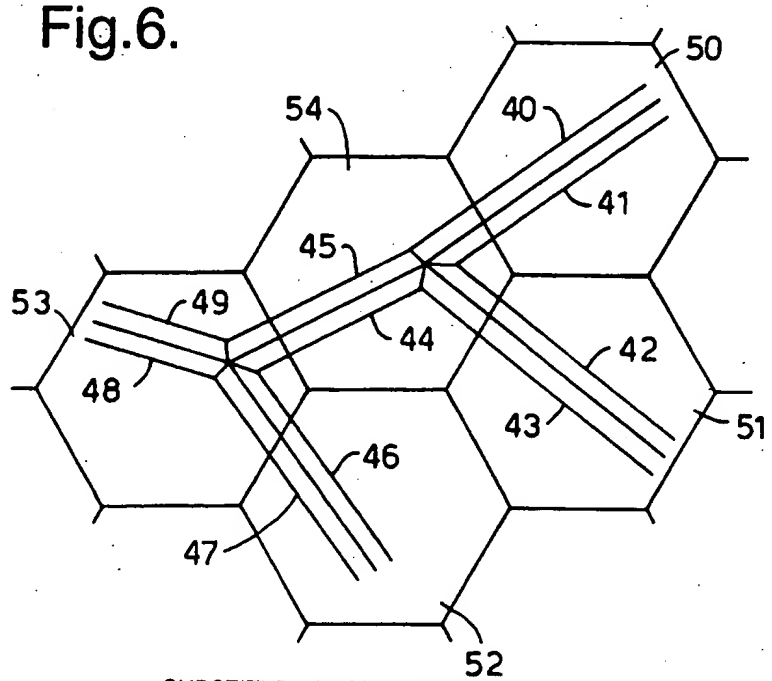


Fig.6.



INTERNATIONAL SEARCH REPORT

Internal Application No
PCT/GB 95/02065

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G01S5/14 G01S5/00 H04Q7/38 G08G1/127

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G01S H04Q G08G

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP,A,0 174 540 (GEOSTAR COOPERATION) 19 March 1986	1,7,9, 17,23, 24,30, 47,48
Y	see page 6, line 1 - page 7, line 24; figure 1	2-6,8, 10-16, 18-22, 25-29, 31,32, 38-46

Y	GB,A,2 176 964 (STC PLC) 7 January 1987 see page 1, line 63 - page 2, line 5; figure 1	8

	-/--	

[X] Further documents are listed in the continuation of box C. [X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
11 December 1995	15. 01. 96
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Haffner, R

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	EP,A,0 379 198 (SHARP KABUSHIKI KAISHA) 25 July 1990 see page 3, line 45 - page 6, line 17; figures 1,2	2-6, 18-22 33-37
Y	EP,A,0 604 404 (CATERPILLAR INC) 29 June 1994 see page 9, line 58 - page 10, line 23 see page 11, line 13 - line 45 see page 12, line 20 - line 24 see page 18, line 16 - line 42; figures 1,3-5,7-10	10-15, 25-29, 31,32, 38-46
Y	EP,A,0 601 293 (MOTOROLA INC.) 15 June 1994 see column 8, line 14 - column 10, line 2 see column 11, line 10 - column 12, line 18; figures 1-4	16,29
A	GB,A,2 264 837 (KOKUSAI DENSHIN DENWA COMPANY LIMITED) 8 September 1993 see page 5, line 14 - page 9, line 12 see page 19, line 4 - page 23, line 15; figures 1-5	1,17, 30-32, 47,48
A	GB,A,2 271 486 (MOTOROLA LIM) 13 April 1994 see page 8, line 11 - line 31 see page 10, line 22 - page 11, line 33; figures 1,2	1,17, 30-32, 47,48

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/GB 95/02065

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-174540	19-03-86	US-A-	4744083	10-05-88
		AU-B-	578970	10-11-88
		AU-B-	4680385	27-03-86
		CA-A-	1249033	17-01-89
		JP-A-	61112439	30-05-86
GB-A-2176964	07-01-87	NONE		
EP-A-0379198	25-07-90	JP-A-	2189488	25-07-90
		JP-A-	2206900	16-08-90
		AU-B-	614893	12-09-91
		AU-B-	4799790	26-07-90
		DE-D-	69021900	05-10-95
		US-A-	5025261	18-06-91
EP-A-0604404	29-06-94	WO-A-	9109375	27-06-91
		EP-A-	0608005	27-07-94
		EP-A-	0679903	02-11-95
		EP-A-	0679975	02-11-95
		EP-A-	0679973	02-11-95
		EP-A-	0679904	02-11-95
		EP-A-	0679974	02-11-95
		EP-A-	0679976	02-11-95
		AU-B-	642638	28-10-93
		AU-B-	7749194	05-01-95
		EP-A-	0507845	14-10-92
		WO-A-	9109275	27-06-91
EP-A-0601293	15-06-94	CA-A-	2106534	08-06-94
		JP-A-	6281720	07-10-94
GB-A-2264837	08-09-93	JP-A-	5232210	07-09-93
GB-A-2271486	13-04-94	NONE		

L2: Entry 4 of 7 File: JPAB Jan 16, 1998

PUB-NO: JP410013946A
DOCUMENT-IDENTIFIER: JP 10013946 A
TITLE: PORTABLE INFORMATION TERMINAL

PUBN-DATE: January 16, 1998

INVENTOR-INFORMATION:

NAME	COUNTRY
FURUYA, MASATOSHI	
KURISU, HIROMITSU	
FUKUMOTO, YASUSHI	

ASSIGNEE-INFORMATION:

NAME	COUNTRY
HITACHI LTD	

APPL-NO: JP08163985
APPL-DATE: June 25, 1996

INT-CL (IPC): H04Q 7/38; G01S 5/14; H04B 7/26

ABSTRACT:

PROBLEM TO BE SOLVED: To improve operational efficiency by displaying information of a communicable or incommunicable position with a map.

SOLUTION: A portable information terminal 100 has a connection interface 300 such as a modem card, etc., to connect a portable telephone 400 and the terminal 100. The terminal 100 has a structure to activate various features which are attached to the terminal 100 by touching a display screen 200 with a pointing device 600, a finger, etc. Also the terminal 100 includes a GPS card 500 which receives an electric wave from a satellite and detects its current location and recognizes the current location of a user of the terminal 100. The current location is inputted, and the information which shows whether or not the current location is a communication position is registered as geographical data attribute information of the current location to geographical database in the terminal 100 and shown, with the symbol overlapped, which shows whether or not the current location is a communicable position within a prescribed range on the map, including the position where the attribute information is registered.

(19)日本国特許庁（ＪＰ）　(12) 公開特許公報（A）　(11)特許出願公開番号

特開平10－13946

(43)公開日　平成10年(1998)１月16日

(51)Int.Cl.⁶		識別記号	庁内整理番号	ＦＩ			技術表示箇所
Ｈ０４Ｑ	7/38			Ｈ０４Ｂ	7/26	１０９Ｔ	
Ｇ０１Ｓ	5/14			Ｇ０１Ｓ	5/14		
Ｈ０４Ｂ	7/26			Ｈ０４Ｂ	7/26	Ｍ	
						１０９Ｍ	

審査請求　未請求　請求項の数７　ＯＬ　（全 11 頁）

(21)出願番号　特願平8－163985

(22)出願日　平成８年(1996)６月25日

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地

(72)発明者　古谷　雅年
神奈川県川崎市麻生区王禅寺1099番地　株式会社日立製作所システム開発研究所内

(72)発明者　栗栖　宏充
神奈川県川崎市麻生区王禅寺1099番地　株式会社日立製作所システム開発研究所内

(72)発明者　福本　恭
神奈川県川崎市麻生区王禅寺1099番地　株式会社日立製作所システム開発研究所内

(74)代理人　弁理士　秋田　収喜

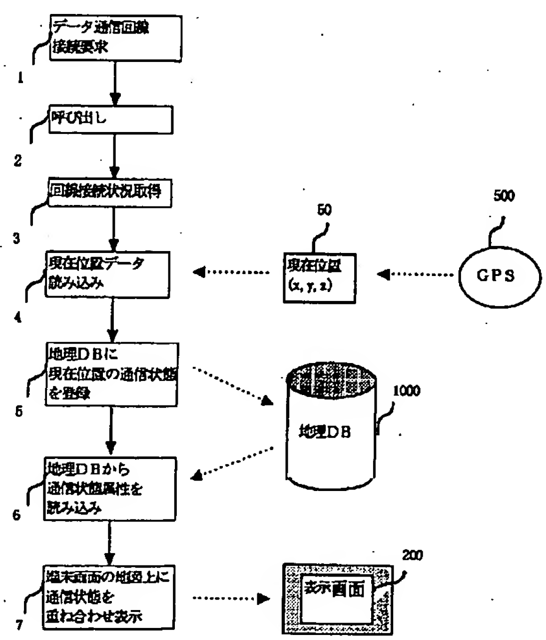
(54)【発明の名称】　携帯情報端末

(57)【要約】

【課題】　携帯情報端末を所持し、携帯電話を介して、他とのデータ通信を行いながら作業する者のために、通信可能または不可能な場所の情報を地図と共に表示する機能を有する携帯情報端末を提供すること。

【解決手段】　携帯電話機を介してデータ通信を行う携帯情報端末において、現在位置を入力し、その現在位置が通信可能位置か否かを示す情報を現在位置の地理データの属性情報として携帯情報端末内の地理データベースに登録し、この属性情報が登録されている位置を含む地図上の所定範囲内に通信可能位置か否かを示す記号を重ね合わせ表示する。

図２



【特許請求の範囲】

【請求項１】　携帯電話機を介してデータ通信を行う通信手段を備えた携帯情報端末であって、
所定区域の地理データを格納した地理データベースと、この地理データベースから地理データを読出し、その地理データに対応した地図を画面上に表示する表示手段と、現在位置のデータを入力する位置データ入力手段と、この位置データ入力手段から入力された現在位置が通信可能位置か否かを示す情報を現在位置の地理データの属性情報として前記地理データベースに登録する属性情報登録手段と、前記属性情報が登録されている位置を含む地図上の所定範囲内に通信可能位置か否かを示す記号を重ね合わせ表示する通信状態表示手段とを備えることを特徴とする携帯情報端末。

【請求項２】　携帯電話機の回線接続後に、その回線接続の中継基地局となった基地局の回線利用状況の情報を携帯電話機の事業体から取得し、その回線利用状況の情報を前記属性情報登録手段に入力し、現在位置の地理データの属性情報として前記地理データベースに登録させる利用状況取得手段とをさらに備え、通信可能位置か否かを示す記号と共に、回線の利用状況を記号または数値で表示することを特徴とする請求項１記載の携帯情報端末。

【請求項３】　前記通信可能位置か否かを示す記号は、属性情報が登録されている地図上の位置、あるいはその位置から所定範囲内の地図上、または所定範囲内に存在する建物等のオブジェクト上に表示することを特徴とする請求項１または２記載の携帯情報端末。

【請求項４】　前記位置データは、端末本体に装着したＧＰＳカードから入力することを特徴とする請求項１ないし３記載のいずれかの携帯情報端末。

【請求項５】　前記位置データは、携帯電話機の回線接続後に、その回線接続の中継基地局となった基地局の位置データを携帯電話機の事業体から取得して入力することを特徴とする請求項１ないし３記載のいずれかの携帯情報端末。

【請求項６】　携帯電話機を介してデータ通信を行う通信手段を備えた携帯情報端末であって、
所定区域の地理データを格納した地理データベースと、この地理データベースから地理データを読出し、その地理データに対応した地図を画面上に表示する表示手段と、携帯電話機の回線接続後に、その回線接続の中継基地局となった基地局の位置データと回線利用状況の情報とを携帯電話機の事業体から取得する基地局情報取得手段と、この基地局情報取得手段が取得した基地局の回線利用状況の情報を当該基地局の位置の地理データの属性情報として前記地理データベースに登録する属性情報登録手段と、前記属性情報が登録されている位置を含む地図上の所定範囲内に基地局の回線利用状況を記号または数値で重ね合わせ表示する通信状態表示手段とを備えることを特徴とする携帯情報端末。

【請求項７】　携帯電話機を介してデータ通信を行う通信手段を備えた携帯情報端末であって、
所定区域の地理データを格納した地理データベースと、この地理データベースから地理データを読出し、その地理データに対応した地図を画面上に表示する表示手段と、現在位置のデータを入力する位置データ入力手段と、この位置データ入力手段から入力された現在位置データを通信相手に送信すると共に、通信相手の現在位置データを通信相手から受信する位置データ送受手段と、前記位置データ入力手段から入力された現在位置および通信相手から受信した通信相手の現在位置に対応する地図上の位置に、それぞれの固有の記号を表示する相対位置表示手段とを備えることを特徴とする携帯情報端末。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、携帯電話機との接続インタフェースを有し、携帯電話機を介してデータ通信を行う携帯情報端末に関するものである。

【０００２】

【従来の技術】ディジタル携帯電話機やＰＨＳなど（以下、単に携帯電話機という）を介してデータ通信を行う携帯情報端末の例として、例えば、「Ｍｏｂｉｌｅ　ＰＣ創刊号」（１９９５年１１月号、Ｐ７６～Ｐ７７、以下文献１という）には、携帯情報端末のカードスロットにモデムカードを挿入し、モデムとＰＨＳとをケーブルで接続してデータ通信を行うものが記載されている。

【０００３】また、携帯電話機の接続状況を表示する例として、「日経トレンディＮｏ．９９」（１９９５年１０月号、Ｐ４２～Ｐ４４、以下文献２という）には、携帯電話機のディスプレイ上に、「接続できない」，「圏外」などの文字を表示するなどして、通信不能状態であることを利用者に知らせるものがある。

【０００４】また、携帯情報端末で地理データを活用する例として、例えば、上記文献１のＰ２３には、ペン入力パーソナルコンピュータ（以下、ペンＰＣ）にホストコンピュータのデータベースの情報をダウンロードし、ペンＰＣを現場に携帯し、地図や設備図、設計図、及び、その属性情報を現場で活用する例が示されている。このとき、携帯情報端末にＧＰＳカードを組み込んで地図上に現在位置を表示することもある。

【０００５】

【発明が解決しようとする課題】ところで、携帯電話機は、基地局との間で無線通信を行うが、通信不能になる例として、上記文献２のＰ１０～Ｐ３１には、（１）基地局が設置されてなく通信不能である、（２）屋内や地下空間、建物の陰などで通信不能になる、（３）地震などの災害発生時に基地局がダウンし通信不能になる、（４）基地局の交換能力を超えた発着信があり、通信不能になる、などを挙げている。

【０００６】このような状況下においては、携帯情報端末を所持する作業者が、携帯情報端末を利用して、情報サーバ、または、他の作業者との間でデータ通信を行いながら作業する必要がある場合、どこの地点が通信可能なのか、あるいは、どこの地点が通信不能なのかを知ることができれば、通信不能になった場合でも、通信可能、または、可能性の高い場所へ移動し易くなる。また、通信可能な場所だけを移動できるようになるので、現場を転々として、データ通信を行いながら作業する必要がある場合、作業効率が良くなる。

【０００７】一方、携帯情報端末を所持する者の中には、現場周辺の土地勘が少ないこともあるため、場所に関する情報は地図とともに提供すると便利である。

【０００８】しかしながら、従来の携帯情報端末には、このような機能が組み込まれていないため、通信可能な場所を探しながらデータ通信を行う必要があるため、作業効率が悪いという問題がある。

【０００９】本発明の目的は、携帯情報端末を所持し、携帯電話機を介して、他とのデータ通信を行いながら作業する者のために、通信可能または不可能な場所の情報を地図と共に表示し、作業効率の向上に貢献することができる携帯情報端末を提供することにある。

【００１０】

【課題を解決するための手段】上記目的を達成するために、本発明は、地理データベースと、地理データに対応した地図を画面上に表示する表示手段と、携帯電話機を介してデータ通信を行う通信手段と、現在位置データを入力する位置データ入力手段と、この位置データ入力手段から入力された現在位置が通信可能位置か否かを示す情報を現在位置の地理データの属性情報として前記地理データベースに登録する属性情報登録手段と、前記属性情報が登録されている位置を含む地図上の所定範囲内に通信可能位置か否かを示す記号を重ね合わせ表示する通信状態表示手段とを備えることを特徴とする。

【００１１】ここで、通信可能位置か否かを示す記号は、属性情報が登録されている地図上の位置、あるいはその位置から所定範囲内の地図上、または所定範囲内に存在する建物等のオブジェクト上に表示する。

【００１２】また、位置データは、ＧＰＳカードを組み込んで入力する方法、あるいは地番等を手入力で入力する方法を用いることができる。

【００１３】また、現在位置が通信可能位置であるか否かは、データ通信回線確立操作を（相手呼出し操作）を行った時に、相手が応答したか否かによって判定することができる。

【００１４】一方、位置データをＧＰＳカードまたは手入力で入力する代わりに、携帯電話機の回線を接続しようとしたとき、回線の接続ができたならば、携帯電話機を運用している事業体より携帯電話機の基地局に関する位置データを取得するようにし、この基地局の位置の地理データに通信可能であることを示す属性情報を登録するようにしてもよい。

【００１５】この場合、基地局の回線利用状況をも取得し、回線の混雑具合等を記号や数値で表示するようにしてもよい。

【００１６】

【発明の実施の形態】以下、本発明の実施の形態を図面に基づき詳細に説明する。

【００１７】図1は、本発明を適用した携帯情報端末の実施の形態を示す外観構成図、図２は本発明の携帯情報端末において地図上の各位置における通信状態を登録する処理及び通信状態を表示する処理の概要を示すフローチャートである。

【００１８】この実施携帯における携帯情報端末１００は、携帯電話機４００と携帯情報端末１００を接続するために、モデムカードなどの接続インタフェース３００を有している。携帯情報端末１００は、ペンなどのポインティング装置６００、または指などで表示画面２００に触ることにより、携帯情報端末１００に設けられた各種の機能を起動するように構成されている。また、人工衛星から電波を受信して現在位置を検出するＧＰＳカード５００を組み込み、携帯情報端末１００の利用者の現在位置を認識できるようになっている。

【００１９】なお、携帯電話機４００の機能を全て携帯情報端末１００内に組み込んだ構成で実現する場合がある。

【００２０】携帯情報端末１００は、所定区域の地理データを格納した地理データベースと、この地理データベースから地理データを読出し、その地理データに対応した地図を表示画面２００上に表示する表示処理手段と、表示画面２００上で指またはポインティング装置６００を用いて現在位置のデータを入力する位置データ入力手段と、この位置データ入力手段から入力された現在位置が通信可能位置か否かを示す情報を現在位置の地理データの属性情報として前記地理データベースに登録する属性情報登録処理手段と、前記属性情報が登録されている位置を含む地図上の所定範囲内に通信可能位置か否かを示す記号を重ね合わせ表示する通信状態表示処理手段とを内部機能として備えている。

【００２１】携帯情報端末１００の表示画面２００には、例えば図3に示すような所定区域の地図２０１が表示されるようになっている。この地図２０１は、地図データベースに格納された地理データのうち利用者が指定した区域の地理データを可視化したものである。

【００２２】この場合、表示させる地図の区域の指定方法としては、表示画面２００上で区域指定を行う方法と、ＧＰＳカード５００が検出した現在位置を中心とする所定範囲の区域を指定する方法があるが、いずれを用いてもよい。

【００２３】以下、図２のフローチャートを参照して携

帯情報端末１００の動作を説明する。

【００２４】まず、携帯情報端末１００は、相手接続先を指定したデータ通信回線接続要求があったならば（ステップ１）、携帯電話機４００によって指定された接続先の呼び出しを行う（ステップ２）。そして、その呼び出しにより回線接続ができたかどうかの接続状況を取得する（ステップ３）。

【００２５】次に、ＧＰＳ５００が検出している現在位置データ５０を読み込み（ステップ４）、地理データベース１０００に、現在位置の属性情報として通信可能であったかどうかの情報を登録する（ステップ５）。なお、現在位置データ５０は、ｘ，ｙ，ｚの３次元データである。

【００２６】この場合、現在位置データ５０に基づいて、地理データベース１０００から建物などのオブジェクトを検索し、このオブジェクトの属性情報として通信可能であったかどうかの情報を地理データベース１０００に登録してもよい。

【００２７】次に、地理データベース１０００から現在位置、または利用者が指定したオベジェクトに対する通信状態に関する属性情報を読み込み（ステップ６）、図３に示したように携帯情報端末１００の表示画面２００に、地図２０１と通信の状態を表す記号２０２，２０３を重ね合わせて表示する（ステップ７）。

【００２８】この場合、通信可能状態２０２、または通信不能状態２０３の一方だけを表示することもある。

【００２９】さらに、現在位置を表す記号２０４も合わせて表示することもある。現在位置を表す記号２０４をも表示させた場合、携帯情報端末１００の利用者の現在位置と通信可能地域との位置関係とが明瞭になり、携帯情報端末１００の利用者は通信可能地域へ簡単に移動することができるようになる。

【００３０】ここで、ＧＰＳ５００を搭載していない携帯情報端末１００を利用している場合には、図４のフローチャートに示すように、現在位置を利用者に入力してもらうように表示画面２００に案内メッセージを表示し（ステップ１１）、利用者に表示画面２００に表示された地図上の現在位置に相当する位置をポインティング装置ペン６００で指定させることで、現在位置データ５０を入力させ（ステップ１２）、現在位置データ５０を読み込むようにしてもよい。

【００３１】また、地図上の現在位置に相当する位置をポインティング装置ペン６００で指定させる代わりに、地番等のデータを入力するようにしてもよい。

【００３２】ここで、図２の処理手順は、データ通信回線接続要求があったときに通信状態の登録及び通信状態の表示の一連の処理を行うものであるが、通信状態の登録までのステップ５までを行い、処理を一旦終える場合もある。

【００３３】また、図５のフローチャートに示すように、携帯情報端末１００の利用者が通信状態の確認要求をしたとき（ステップ２１）、利用者が指定した位置、またはその位置を含む所定範囲の地理データと通信状態情報（属性情報）を地理データベース１０００から読出し（ステップ６）、表示画面２００に地図と共に、図３のように通信状態の可否を示す記号を重ね合わせ表示するようにしてもよい。

【００３４】通信状態の確認要求方法としては、携帯情報端末１００のメニューやアイコンやボタンを使い、通信状態の属性情報の表示モードと非表示モードを切り替える方法や、通信確認要求アイコンをポインティング装置６００でクリックする方法などがある。

【００３５】図６は、通信状態を登録する機能、及び通信状態を表示する機能の他の実施形態を示すフローチャートである。

【００３６】利用者が携帯電話機４００を使って発信し（ステップ２）、相手接続先と回線が接続されたならば、携帯電話機４００の運用を行っている事業体は、回線交換局または制御局８００にて発信元の位置検索を行う（ステップ３１）。すなわち、携帯電話機４００の回線接続を中継した基地局の位置データを検索する。そして、その検索した基地局の位置データ６０を発信元の携帯電話機４００に送信する（ステップ３２）。

【００３７】携帯情報端末１００は、携帯電話機４００を介して位置データ６０を取り込む（ステップ３４）。そして、その取り込んだ位置データに基づき、基地局の位置とその属性情報として通信可能であることを示す情報を地理データベース１０００に登録する（ステップ３５）。

【００３８】このように、携帯電話機４００の運用を行っている事業体から回線接続を中継した基地局の位置データを取得し、この位置データで示される位置の地理データに通信可否を示す情報を属性情報として登録することにより、ＧＰＳカード５００や利用者の入力による現在位置の特定がなくても、図７に示すように、携帯情報端末１００の表示画面２００の地図２０１上に、基地局位置２１１を中心とした通信可能範囲２１２を表示することができる。

【００３９】さらに、ＧＰＳカード５００などで特定された現在位置５０と組み合わせ、現在位置を表す記号２０４を表示すれば、現在位置と通信可能範囲との位置関係が地図上でさらに明瞭になる。

【００４０】なお、図７においては、発信元だけでなく、その近傍の複数の基地局の通信可能範囲を同時に表示しているが、発信元の基地局のみの通信可能範囲を表示するようにしてもよい。

【００４１】図８は、本発明の通信状態を登録する機能及び通信状態を表示する機能の他の実施形態を示すフローチャートである。

【００４２】まず、利用者が携帯電話機４００を使って

発信し（ステップ２）、相手接続先と回線が接続されたならば、携帯電話機４００をサービスしている事業体の回線制御局８００は、その回線接続を中継した基地局を含む所定範囲内の複数の基地局の位置データ２０５と回線利用状況情報（回線利用数などの情報）２０６とを基地局データベース２０００から読み込み（ステップ４１）、発信元の携帯電話機４００に転送する（ステップ４２）。

【００４３】携帯情報端末１００は、携帯電話機４００を介して基地局の位置データ２０５および回線利用数情報２０６を読み込み（ステップ４４）、さらに、携帯情報端末１００の内部時計から現在時刻情報２０７を読み込み（ステップ４５）、地理データベース１０００の各基地局の位置データ２０５で示される位置の地理データに回線利用状況情報２０６及び現在時刻情報を属性情報として登録する（ステップ４６）。

【００４４】そして、図9に示すように、携帯情報端末の表示画面２００の地図２０１上にかく基地局位置２１１を中心とした通信可能範囲２１２を表示し、さらに基地局位置には、回線利用状況を示す数値を表示する。回線利用状況を示す数値は、「０」が回線利用数が全くないことを示し、数値が大きくなるほど、回線利用数が多いことを示している。

【００４５】これによって、どの地域の回線が混雑しているかが地図上で明瞭になる。さらに、表示画面２００の上部に回線利用状況を入手した時刻２２２である「何時何分の状態です」、あるいは、現在時刻と入手した時刻の時間差「何分前の状態です」を表示することにより、いつの情報を表示しているかがわかり、利用者にとって便利である。

【００４６】図１０は、本発明の通信元及び通信先の位置を表示する機能の他の実施形態を示すフローチャートである。

【００４７】まず、通信元の携帯電話機４００の利用者が通信相手を指定して接続要求を行うと（ステップ１）、通信元の携帯電話機４００は通信相手の呼出しを行う（ステップ２）。

【００４８】これによって、通信元と通信相手との回線が確立したならば（ステップ６３）、通信元及び通信相手はそれぞれ現在位置データ５０、６０を読み込み（ステップ５４，６４）、お互いに相手に現在位置情報を送信し（ステップ５５、６５）、相手の現在位置データを取得する（ステップ５６、６６）。

【００４９】次に、通信元では、図１１に示すように、携帯情報端末１００の表示画面２００に地図２０１と通信元及び通信相手の位置を表す記号２３１，２３２を重ね合わせて表示する（ステップ５７）。

【００５０】通信元では、回線確立中は、ステップ５４から５７までの処理を繰り返す。通信相手側も同様である。

【００５１】これによって、双方の現在位置を地図上に表示でき、互いに関連する作業を行っているような場合に、相手の位置が明瞭になり、物的または人的な支援などを効果的に行うことが可能になる。

【００５２】なお、本発明の実施の形態では、地図に通信状態に関する属性情報のみを表示するものとして説明したが、地図上の道路の混雑具合や、通行可否状態など、その他の属性情報を同時に重ね合わせ表示してもよい。地図上の道路の混雑具合などの情報は、交通情報を提供している機関や団体から容易に取得することができる。

【００５３】

【発明の効果】以上のように本発明によれば、携帯電話機を介してデータ通信を行う携帯情報端末において、現在位置を入力し、その現在位置が通信可能位置か否かを示す情報を現在位置の地理データの属性情報として携帯情報端末内の地理データベースに登録し、この属性情報が登録されている位置を含む地図上の所定範囲内に通信可能位置か否かを示す記号を重ね合わせ表示するようにしたため、携帯情報端末を所持し、携帯電話機を介して、他とのデータ通信を行いながら作業する場合などにおいて、どの地域が通信可能な地域であるのか、あるいは通信不可能な地域であるのかを極めて容易に認識することができるため、作業効率を向上させることができる。

【００５４】また、位置データをＧＰＳカードから取得するようにした場合、位置データを入力する手間がなくなり、どの地域が通信可能な地域であるのか、あるいは通信不可能な地域であるのかを極めて容易に認識させることができる。

【００５５】さらに、通信可否の記号と共に、現在位置を示す記号を同時画面内に表示するようにした場合、現在位置と通信可能地域との位置関係が明瞭になり、通信可能地域への移動、あるいは物的、人的支援が容易になる。

【００５６】また、回線の接続状況を示す情報を携帯電話機運用事業体から取得して表示するようにした場合、回線が混雑していない地域に移動することにより、回線が空き状態になるまで待機することなく直ちに通信が可能になり、緊急時のデータ通信を遅滞なく行うことが可能になる。

【００５７】さらに、通信相手相互間で現在位置データを送受し、その位置に所定の記号を表示することにより、通信相手の位置関係が明瞭になり、物的、人的支援が容易になるなどの効果がある。

【図面の簡単な説明】

【図１】本発明を適用した携帯情報端末の実施の形態を示す外観構成図である。

【図２】通信状態を登録する機能及び表示する機能の第１の実施の形態を示すフローチャートである。

【図3】地図上に通信可否の状態を表示した例を示す説明図である。

【図4】現在位置データを手入力する場合の手順を示すフローチャートである。

【図5】通信可否の状態を表示させる場合の他の手順を示すフローチャートである。

【図6】通信状態を登録する機能及び表示する機能の第2の実施の形態を示すフローチャートである。

【図7】地図上に通信可否の状態を表示した他の例を示す説明図である。

【図8】通信状態を登録する機能及び表示する機能の第3の実施の形態を示すフローチャートである。

【図9】地図上に通信可否の状態を表示した他の例を示す説明図である。

【図10】通信元と通信相手の位置を表示する機能の処理手順を示すフローチャートである。

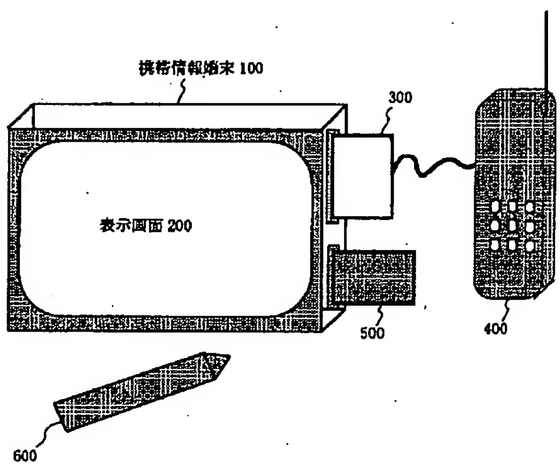
【図11】通信元と通信相手の位置の表示例を示す説明図である。

【符号の説明】

100…携帯情報端末、200…表示画面、300…接続インタフェース、400…携帯電話機、500…GPSカード、600…ポインティング装置、1000…地理データベース、201…地図。

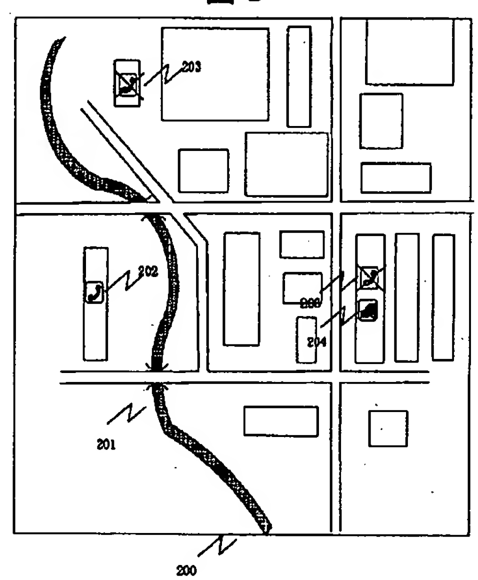
【図1】

図1



【図3】

図3



【図2】

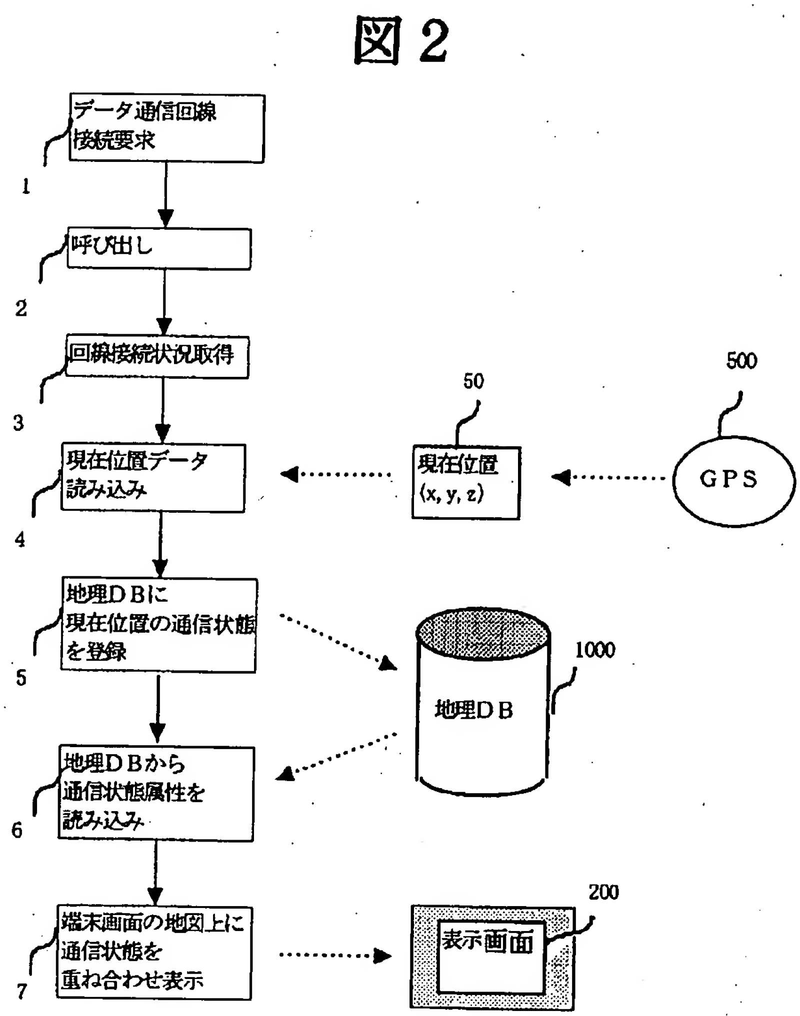


図2
データ通信回線接続要求
1
呼び出し
2
回線接続状況取得
3
現在位置データ読み込み
4
地理ＤＢに現在位置の通信状態を登録
5
地理ＤＢから通信状態属性を読み込み
6
端末画面の地図上に通信状態を重ね合わせ表示
7
50
現在位置
(x, y, z)
500
GPS
地理ＤＢ
1000
表示画面
200

【図4】

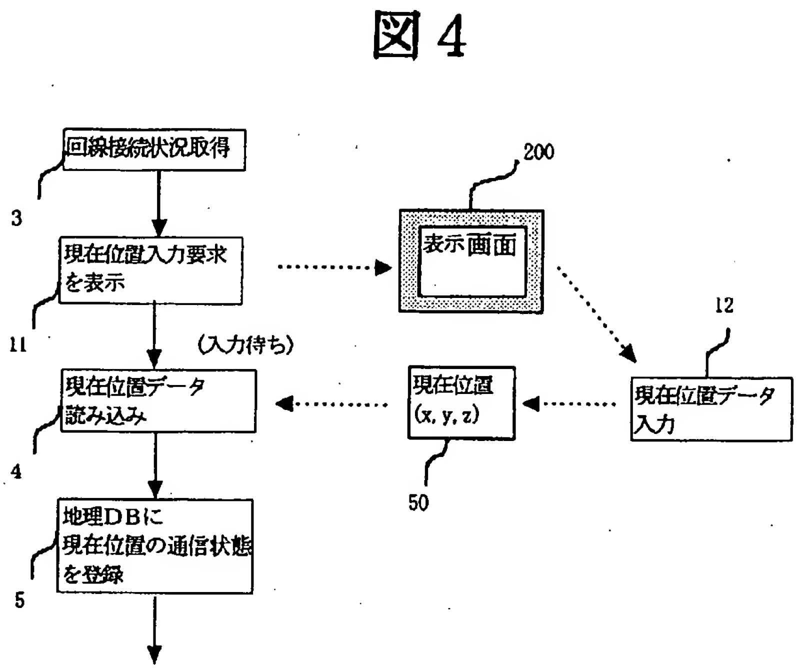


図4
回線接続状況取得
3
現在位置入力要求
を表示
11
(入力待ち)
現在位置データ
読み込み
4
地理DBに
現在位置の通信状態
を登録
5
200
表示画面
12
現在位置データ
入力
現在位置
(x, y, z)
50

【図5】

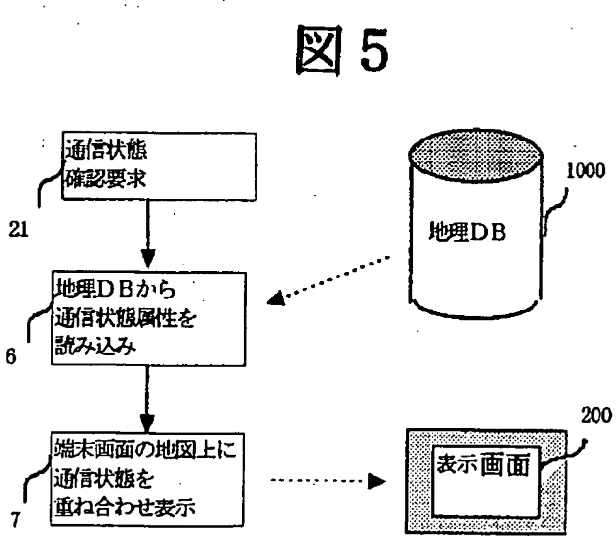


図5
通信状態
確認要求
21
地理DBから
通信状態属性を
読み込み
6
端末画面の地図上に
通信状態を
重ね合わせ表示
7
地理DB
1000
表示画面
200

【図7】

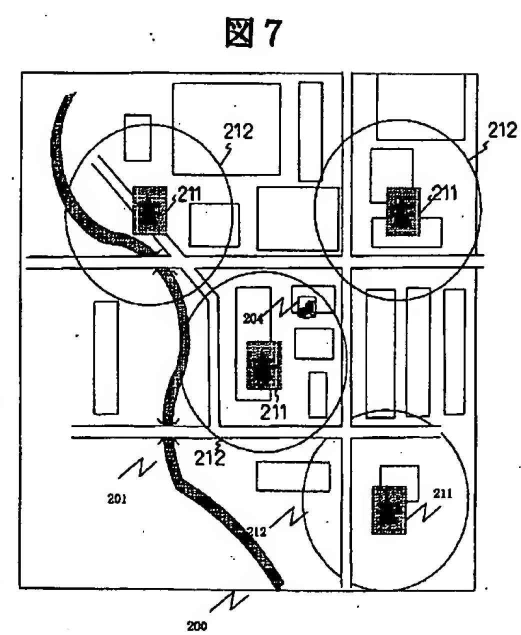


図7
212
211
204
201
200

【図6】

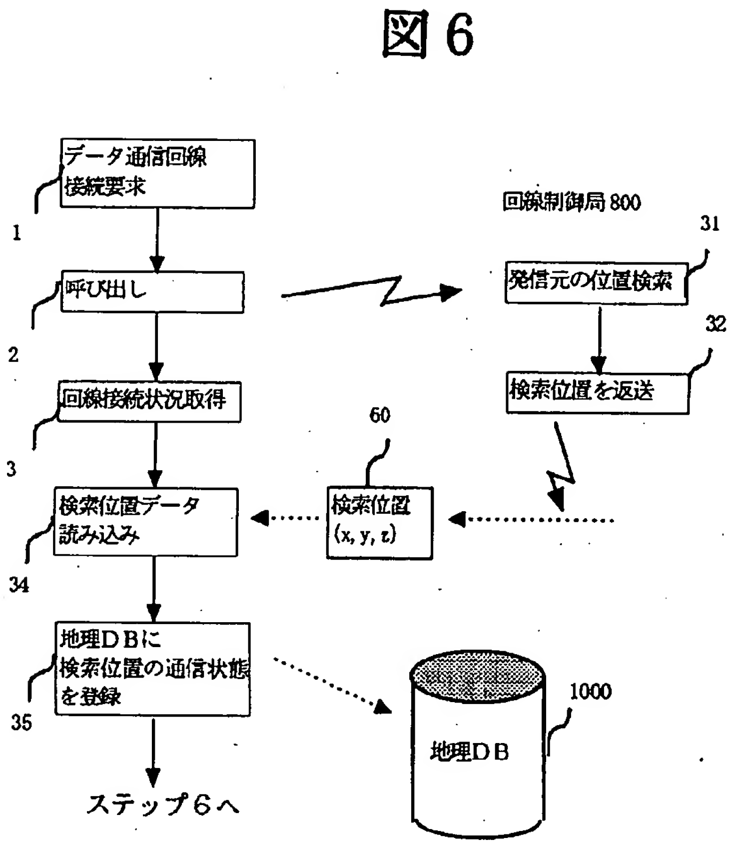
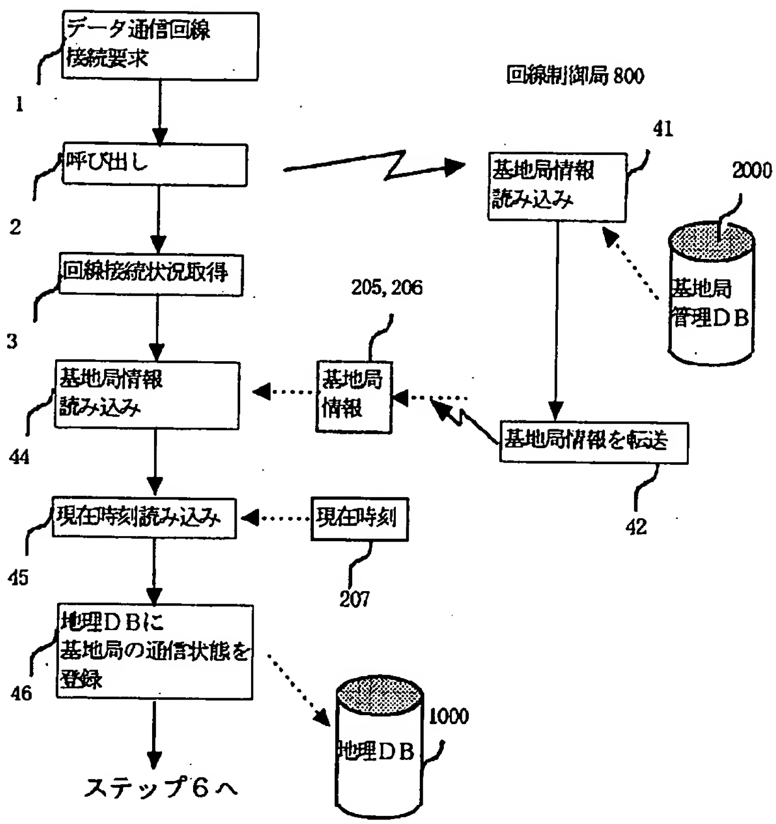


図6
データ通信回線
接続要求
1
呼び出し
2
回線接続状況取得
3
検索位置データ
読み込み
34
地理ＤＢに
検索位置の通信状態
を登録
35
ステップ６へ
回線制御局 800
31
発信元の位置検索
32
検索位置を返送
60
検索位置
(x, y, z)
1000
地理ＤＢ

【図8】

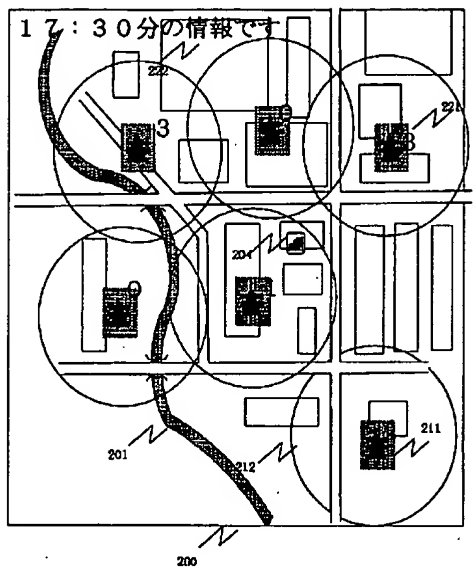
図8



データ通信回線接続要求
1
呼び出し
2
回線接続状況取得
3
基地局情報読み込み
44
現在時刻読み込み
45
地理DBに基地局の通信状態を登録
46
ステップ6へ
回線制御局800
基地局情報読み込み
41
2000
基地局管理DB
205, 206
基地局情報
基地局情報を転送
42
現在時刻
207
地理DB
1000

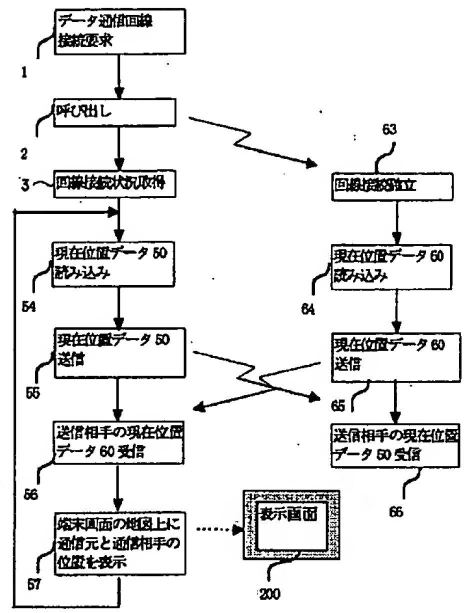
【図9】

図9



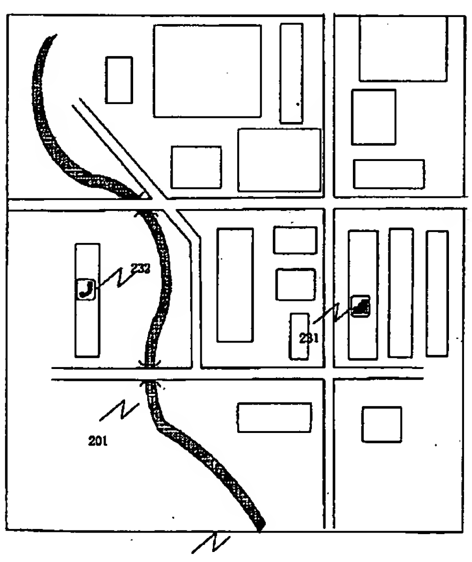
【図10】

図10



【図11】

図11



Join Strategies on KD-Tree Indexed Relations

Masaru KITSUREGAWA, Lilian HARADA, Mikio TAKAGI

Institute of Industrial Science
University of Tokyo
22-1, Roppongi 7, Minato-ku, Tokyo
Japan

ABSTRACT

In this paper we present efficient join algorithms on very large relations indexed by KD-trees. There are previous works proposing the join on multi-attribute clustered relations based on hashing and also on grid-partitioning, whose shortcomings are non-order preservation and low load-factor, respectively. KD-tree indexed relations are characterized by preserving data order and maintaining high load-factors. However, KD-tree indexing has the disavantadge of generating clusters which are overlapped in the join attribute domain, what causes a very high I/O cost for naive join algorithms. Here we analyze strategies to deal with this problem and introduce efficient algorithms to join two non-resident relations indexed by KD-trees.

First we introduce the concept of wave, which is a set of pages that is the object of join processing and that propagates over the relation space in the direction of the join attribute axis. Based on this new concept, we present five join algorithms and also four extended algorithms with a garbage collection mechanism to increase the effective space of the main memory. We extensively evaluate these join algorithms with analytical formulas and simulation results. It is shown that the join of very large relations indexed by KD-trees can be performed with one scan of the relations.

1. Introduction

Recently, the trend of research on very large relations is towards efficient multidimensional clustering of data on secondary memory in order to reduce processing cost, principally I/O cost.

The multidimensional clustering methods can be classified into two broad categories : the fixed methods which partition the data space at fixed places, and the adaptable methods which partition the data space according to the data distribution. Some examples of the fixed methods are the Grid File[10], DYOP[13], PLOP[9], Colored Binary Trie[16], Multidimensional Digital Hashing[12], Multidimensional Extendible Hashing[11,7,8] and some examples of the adaptable methods are the KD-Tree[1,2], Extended KD-Tree[3], KDB-Tree[15], GKD-Tree[5] and R-Tree[6]. However, almost all the analysis based on these multidimensional clustering methods are restricted to the selection and aggregation operations.

Concerning the join operation, some of the recent works are based on Predicate-Tree[18,4], Superjoin Algorithm[17] and DYOP[14]. The first two works use clustering based on multidimensional hashing and the last, on grid-partitioning, all of them being multi-attribute clustering mechanisms based on the fixed method. Fixed clustering methods are proper for join operations because all tuples with a discriminator attribute value share the same cluster and, the join of two relations is thus reduced to the join of their correspondent clusters. However, there are some shortcomings : the methods based on hashing do not preserve data order, which makes them inefficient for range-queries, and those based on grid-partitions have low load-factors, which increases the I/O cost.

The KD-tree indexing is one of the most known multidimensional clustering mechanisms based on the adaptable method, with the advantages of order-preservation and high load-factor but the disadvantage that tuples with a given discriminator attribute value are not in a single cluster. A naive join algorithm on very large relations indexed by KD-trees presents a high I/O cost and until now there is no work investigating join algorithms using KD-trees.

In this paper we propose join algorithms on KD-tree indexed relations. The new join algorithms are based on a new concept called wave. Wave is a set of pages that is the object of joining and that propagates over the relation space in the direction of the join attribute axis. Here we propose four basic join algorithms that determine the wave from one of the relations, and one algorithm that determines the wave from both relations. After describing these algorithms, we extensively analyze them with analytical formulas and simulation results. Then we introduce a garbage collection mechanism that discards the unnecessary data loaded in the main memory and extends the previous basic algorithms with an efficient memory management. We extensively evaluate these algorithms with analytical formulas and simulation results and show that the proposed algorithms perform the join of very large relations with one scan.

Section 2 provides some background to the proposed algorithms, describing some notation and assumptions used in our analysis. Section 3 describes how, using the KD-tree indexing mechanism, the join can be performed on very large relations. It presents five join algorithms for KD-tree indexed relations, their analytical and simulation results. Section 4 describes how the dynamic removal of unnecessary tuples from the main memory increases its effective space and improves the previous algorithms. If there is a suitable memory management, each relation is read into the main memory once. We present four extended versions of the join algorithms presented in section 3, their cost formulas and simulation results. Section 5 concludes the paper.

2. Notations and Assumptions

Let R be a relation having k attributes $A_1,...,A_k$ and composed of tuples $t = (a_1,...,a_k)$.

D is the base space of relation R and denotes the Cartesian product of the domains of attributes referred to by the relation, i.e.,

$$D = \prod_{i=1}^{k} [\, MIN_i \,, MAX_i \,]$$

For a given relation, the KD-tree method first divides D into two subspaces, using the A_1 value of the tuple so that it evenly divides the whole set of tuples. Here A_1 is called discriminator attribute. This step is then recursively applied to the subspaces with the discriminator attributes changed in the cyclic order as A_2, A_3, ...,

CH2695-5/89/0000/0085$01.00

A_k, A_1, ... until each produced subspace can be contained in a disk page. The KD-tree is gradually constructed while dividing D. When a subspace is divided, a new node, which contains the discriminator attribute, its value at the recursive step, as well as a right and left pointers to the two new subspaces, is added to the KD-tree.

Denoting the $|R|$ pages generated by dividing relation R as p_j $(1 \le j \le |R|)$, the space P_j of pages p_j is represented by :

$$P_j = \prod_{i=1}^{k} [\alpha_{ij}, \beta_{ij}] \quad (\alpha_{ij} < \beta_{ij}, \alpha_{ij}, \beta_{ij} \in [MIN_i, MAX_i] \text{ for } \forall i).$$

We define "cluster of an attribute value a_i" the set of pages p_j $(1 \le j \le cs)$ whose space contains the attribute value a_i, i.e.,

$$\text{cluster of } a_i = \{ p_j \mid a_i \in [\alpha_{ij}, \beta_{ij}] \}.$$

When the KD-tree is perfectly balanced, the cluster size cs equals $|R|^{(1-1/k)}$ pages. An example of a cluster of a_1 for a 16-page relation and a 2-dimensional KD-tree is illustrated in Fig. 1.

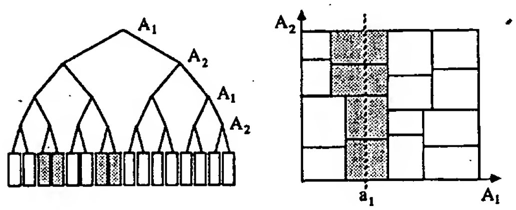


Fig. 1 KD-tree Indexed Relation and a Cluster

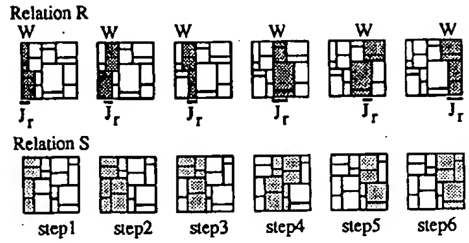
3. Join Algorithms on KD-tree Indexed Relations

In this section we present four algorithms to join two non-resident relations R and S of sizes $|R|$ and $|S|$ pages on attribute A_i. Our analysis is restricted to perfectly balanced KD-trees. Here we consider $|S| \ge |R| >> M$, where M is the main memory size.

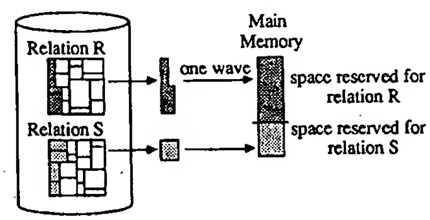
3.1 Description of the Basic Algorithms

A join of two KD-tree indexed relations R and S is illustrated in Fig. 2. In this example, the relations which are composed of 16 pages and indexed by a 2-dimensional KD-tree are joined on attribute A_1. As shown in the figure, the join of R and S is processed in six steps. In each step, four pages of relation R are loaded in the main memory and the range of the join attribute value to process is determined. Here we call this range as "Join Range Jr". The pages of relation S which contain tuples whose join attribute values a_1 are in the join range are loaded in the remaining space of the main memory in a nested-loop way. When the processing in this join range finishes, the two pages of relation R, all of whose tuples have already been processed, are unloaded and then two new pages are loaded. A new join range is determined and the join is processed in the same way as in step 1. As shown in Fig. 2, this procedure is repeated for six steps, when R and S are joined in the entire range of the join attribute A_1.

The set of pages of relation R loaded in the main memory is the unit of processing in each step and because of its shape, we denote it as "Wave W". The average speed of wave propagation over the relation space in the direction of the join attribute axis is given by 1/(number of join steps). In the example of Fig. 2 the wave size is 4 pages, the number of join steps is 6 and the wave propagation speed is 1/6. In this example, the waves of consecutive join steps have 2 pages in common, i.e., 2 pages of a wave in a join step overlap with the wave in the next join step.



(a) Processing Pages of Relations R and S in Each Join Step



(b) Join Processing in Step 1

Fig. 2 Processing Overview of Join Algorithms on KD-Tree Indexed Relations

In the join algorithms presented here, the wave size and the join range are two important parameters. Here we consider that, in a wave, all the tuples whose attribute value is in the join range are processed in each step. Also, in order to hold down the I/O cost of relation S, the wave, in each step, contains at least one cluster of an attribute value in the join range. This implies that the minimum wave size is cs pages and the maximum wave size is (M-1) pages. Also, as shown in Fig. 3, the join range can be determined in two ways : Jr_{in} and Jr_{out}. Considering a wave W propagating in the direction of the join attribute A_i, we call "Wave Rear Line Wr" the minimum join attribute values for the cs pages of the rear of the wave W, and "Wave Front Line Wf" the maximum join attribute values for the cs pages of the front of the wave W. Therefore the two possible join ranges Jr_{in} and Jr_{out} can be expressed as :

$$Jr_{in} = [\max(\alpha_{ij} \in Wr), \min(\beta_{ij} \in Wf)]$$

and

$$Jr_{out} = [\min(\alpha_{ij} \in Wr), \max(\beta_{ij} \in Wf)].$$

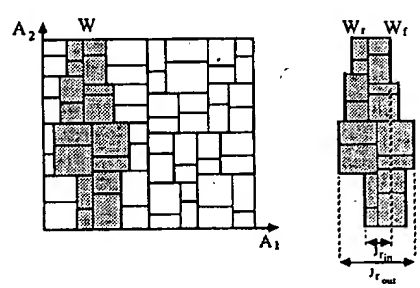


Fig. 3 Waves W and Join Ranges Jr

When the join range is Jr_{in}, waves of consecutive join steps overlap in some pages. On the other hand, when the join range is Jr_{out}, all the pages of waves of consecutive steps are different.

Using the specified wave sizes and join ranges as parameters, we can define four join algorithms as summarized in Table 1.

join range / wave size	Jr_{out} $[\min(\alpha_{ij} \in W_r), \max(\beta_{ij} \in W_r)]$	Jr_{in} $[\max(\alpha_{ij} \in W_r), \min(\beta_{ij} \in W_r)]$
$\lvert R\rvert^{(1-\frac{1}{k})}$	algorithm 1	algorithm 3
M-1	algorithm 2	algorithm 4

Table 1 Wave Size and Join Range for the Basic Algorithms 1, 2, 3 and 4

Following, we detail how the wave W and the join range Jr are determined in each algorithm.

Algorithm 1 :

In this algorithm, the wave is determined as a cluster of relation R and the wave size is cs pages. After loading a wave in the main memory, the join range is determined as Jr_{out}. The pages of relations S whose join attribute value are in the join range are loaded in the remaining (M-cs) pages of the main memory. When processing of the join range Jr_{out} is finished, all the tuples of the wave have already been processed and are thus unloaded. Following, a new wave of cs pages is loaded and the procedure is repeated until the wave propagates over the entire space of relation R. The waves, the corresponding join ranges and their propagation over the relation space are exemplified in Fig. 4. Because the clusters generated by a KD-tree index overlap in the join attribute domain and, in algorithm 1 the join range is Jr_{out}, the waves of consecutive steps are not overlapped but the join ranges are.

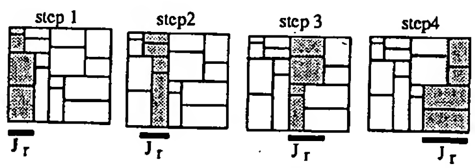


Fig. 4 Waves and Join Ranges for Algorithm 1

Algorithm 2 :

In this algorithm, in order to speed up the wave propagation in comparison with algorithm 1, the wave size is increased to (M-1) pages. Only one page is reserved for relation S. Similarly to algorithm 1, the join range is determined as Jr_{out}, and so waves in consecutive steps are not overlapped while the join ranges are. The waves and corresponding join ranges are exemplified in Fig. 5, for M taken as 7 pages.

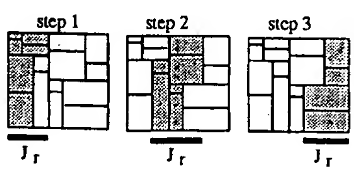


Fig. 5 Waves and Join Ranges for Algorithm 2

Algorithm 3 :

In the same way as in algorithm 1, the wave is determined as one cluster and (M-cs) pages are reserved for relation S. However, in this algorithm, the join range is Jr_{in}, so that waves of consecutive steps overlap and the corresponding join ranges do not. Fig. 6 illustrates the waves and the respective join ranges in algorithm 3.

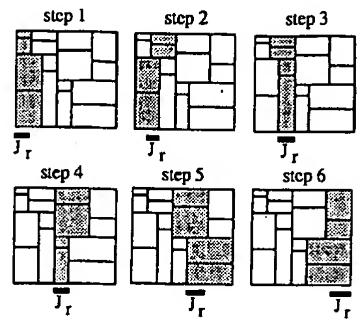


Fig. 6 Waves and Join Ranges for Algorithm 3

Algorithm 4 :

In this algorithm the wave size is (M-1) pages and only one page is reserved for relation S. The join range is determined as Jr_{in}, like in algorithm 3, so that waves in successive steps overlap and the join ranges do not. The wave propagation speed is higher than in algorithm 3. The wave and the join ranges for this algorithm are illustrated by an example in Fig.7.

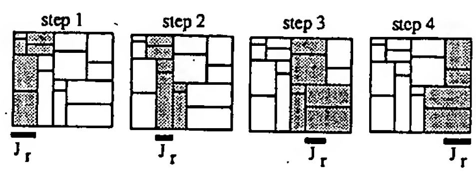


Fig. 7 Waves and Join Ranges for Algorithm 4

In the four algorithms above, the wave and corresponding join ranges are determined from relation R, which is read only once.

Comparing the wave propagation speed for the four algorithms, we have that :

(a) the wave propagation speed for algorithms 2 and 4 is higher than that for 1 and 3, respectively, because the first two have larger wave size (M-1 > cs);

(b) the wave propagation speed for algorithms 1 and 2 is higher than that for 3 and 4, respectively, because the first two have larger join range ($Jr_{out} > Jr_{in}$).

In order to clarify the characteristics of the four basic join algorithms presented above, we present their evaluation results as follows.

3.2 Evaluation for Uniform Data Distribution

3.2.1 Analytical Evaluation

In the previous four algorithms, relation R is read only once and thus, considering $|R| = |S| = N = 2^P$ pages, the number of page

accesses for relation R, π_R, is N pages. Following, we present the expressions which estimate the number of page accesses for relation S, π_S, which is given by :

π_S= (number of join steps) * (pages of relation S within the join range and not in the main memory)..................[1]

For simplification we will use cs to denote the cluster size, which equals $N^{(1-1/k)}$ pages.

Algorithm 1 :

With this algorithm, the number of page accesses of relation S shown in [1] is given by :

π_S= (number of join steps) * {(pages within the join range) - (pages remaining in the main memory)}...............[2]

- the number of join steps is equal to the number of clusters, i.e., N/cs ;
- the number of pages of S within the join range is, on average, 3cs ; and
- the number of pages of S which can remain in the main memory is (M - cs).

Therefore :

$$\pi_S = N\left(4 - \frac{M}{cs}\right)$$

Algorithm 2 :

In this case, only one page is used for relation S, so :

π_S = (number of join ranges) * (pages within the join range).....[3]

- the number of join steps is

$$\frac{N}{M-1};$$

and
- the number of pages of S within the join range is, on average,

$$\left(\left\lceil\frac{M-1}{cs}\right\rceil + \frac{M-1}{cs}\right) cs$$

Therefore :

$$\pi_S = N\left(\left\lceil\frac{M-1}{cs}\right\rceil \frac{cs}{M-1} + 1\right)$$

Algorithm 3 :

In the same way as algorithm 1, π_S can be expressed by [2]. For this case :
- the number of join steps is

$$\sum_{i=1}^{\frac{\ell}{k}-1} 2^{ki} + 2;$$

- the number of pages of S within the join range is, on average, cs ; and
- the number of pages of S saved in the main memory which can be used again in the next join range is

$$M - 1 - cs - 2^{k-1}$$

Therefore :

$$\pi_S = \left(\sum_{i=1}^{\frac{\ell}{k}-1} 2^{ki} + 2\right)(2cs - M + 1 + 2^{k-1}).$$

Algorithm 4 :

In the same way as algorithm 2, π_S can be given by [3]. Here :
- the number of join steps is

$$\frac{N}{(M - 1 - cs) + 2^{k-1}};$$

and
- the number of pages of S within the join range is (M-1) pages.

Therefore :

$$\pi_S = \frac{N(M-1)}{M - cs - 1 + 2^{k-1}}$$

3.2.2 Simulation Results

Following, we present the relative performance of the join algorithms described above, for relations R and S with 64 K pages, 8 tuples/page, 2-dimensional KD-tree and a random data distribution as the one exemplified in Fig. 8.

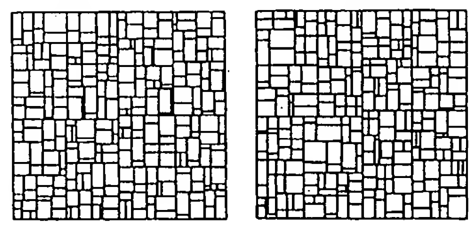


Fig. 8 An Example of Uniform Data Distribution for Relations R and S

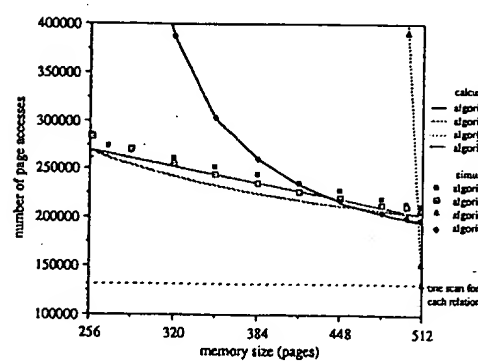


Fig.9 Analytical and Simulation Results of Algorithms 1~4 for Uniformly Distributed Relations

Fig. 9 shows the results of the simulation and analytical evaluation. The vertical axis is the total number of page accesses, π (= $\pi_R + \pi_S$). The horizontal axis is the main memory size M, which is varied from (cs+1) to 2cs pages. We can see that the cost formulas given in section 3.2.1 estimate the number of page accesses quite satisfactorily. Following we compare the algorithm performances for small and large wave sizes.

(a) Small wave size :

From Fig.9 we can observe that, in algorithms 1 and 3, the number of I/O decreases linearly with the memory size M. As shown in the cost formula [1], the number of page accesses for relation S, π_S, depends on the number of join steps and pages of S to read in each step. Given a relation, for algorithms 1 and 3, the number of join steps is constant, and the pages of S to read in each step decreases linearly when M enlarges. The inclination of the curve of algorithm 3 is much greater than that of algorithm 1. As shown by the analytical formulas, this inclination is given by the number of join steps. The number of join steps is much greater in algorithm 3 because the join range is given by Jr_{in}. For small M the I/O cost of algorithm 3 is much higher than that of algorithm 1. On the other hand, when M is enlarged and taken as 2cs pages, in algorithm 3 the I/O cost is reduced to one scan of each relation, while in algorithm 1 this I/O cost is greater.

(b) Large wave size :

Fig. 9 shows that the I/O cost is inversely proportional to M, for algorithms 2 and 4. As given in [1], π_S is proportional to the number of join steps and the pages of S to read. The number of join steps is inversely proportional to the wave size which is (M-1) pages for algorithms 2 and 4. Concerning the inclination of the curves and the I/O cost for small M, the inclination for algorithm 2 is smaller and the performance for small M is better than that for algorithm 4, analogously to algorithm 1 and 3, respectively. However, the difference of algorithms 2 and 4 in comparison with algorithms 1 and 3 is that the I/O cost does not reduce to one scan of the relations, in the range of variation of M shown in Fig. 9.

Observing the four curves we find that the higher the wave propagation speed over the space of relation R, the better the algorithm performs for most range of M. That is, when the main memory space is not enough to maintain the necessary pages of relation S in each step, the performance is better when the number of join steps is smaller. As shown by the formulas, the number of join steps decreases in the sequence of algorithms 3, 4, 1, 2, i.e.,

$$\sum_{i=1}^{k-1} 2^{ki} + 2 > \frac{N}{(M-1-cs)+2^{k-1}} > \frac{N}{cs} > \frac{N}{M-1}$$

for k=2, cs<M<2cs, and so the speed of propagation of the waves increases and the algorithms performance improves in this sequence.

3.3 Algorithm for Non-Uniform Data Distribution

From the analysis of the four basic join algorithms presented above we conclude that :

(1) for small memory sizes, the performance is better when the wave propagation speed is higher.

(2) the only algorithm to reduce the number of page accesses to one scan of the relations, is the one with small wave size and short join range (Jr_{in}) - algorithm 3.

These conclusions and analysis are for relations whose data are uniformly distributed over the relation space and so, for the case in which the wave propagation speed is constant. However, the join range depends on the data distribution and for the case of non-uniformly distributed data, the wave does not propagate with a constant speed.

Following, we introduce a new algorithm for relations with non-uniform data distribution. For this algorithm we use a small wave size and short join range and, using the information of the data distribution for relations R and S, we choose the waves consulting the KD-trees of both relations, aiming at reducing the number of join steps, i.e., maximizing the average wave propagation speed in direction of the join attribute axis.

Algorithm 5 :

In algorithm 5 the wave is taken as a cluster from one relation and (M-cs) pages are reserved for the other relation. The join range is taken as Jr_{in}. In each join step, the wave is taken from one of the relations in order to maximize the wave propagation speed. The waves, their corresponding join ranges and their propagation are illustrated in Fig.10. For the exemplified relations, six join steps are determined with algorithm 3, while algorithm 5 reduces this number to three.

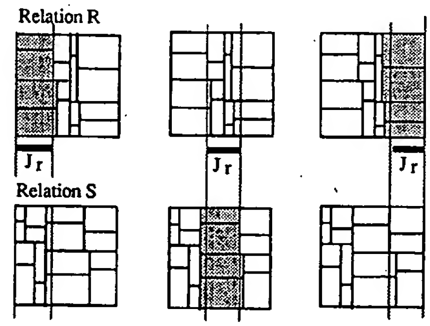


Fig. 10 Waves and Join Ranges for Algorithm 5

3.4 Simulation Results for Non-Uniform Data Distribution

In order to analyze the performance of relations whose data are biased in some regions, we used relations with normal and inverse normal distributions for the simulation. An example of such relations R and S is shown in Fig.11. The sizes of the relation and page used for the simulation are the same as in the uniform data distribution case.

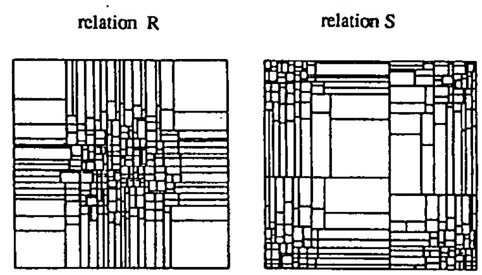


Fig. 11 An Example of Non-Uniform Data Distribution for Relations R and S

The simulation results are shown in Fig.12. As in the case of uniform data distribution, on the first four algorithms the number of page accesses increases in the sequence of algorithms 2, 1, 4, 3, which is the sequence of decreasing the wave propagation speed. Concerning algorithm 5, whose speed of propagation of wave is higher than that of algorithm 3, the performance is improved as expected. Although it is not satisfactory compared to the other algorithms, as shown below in section 4, the extended version of algorithm 5 introduce improvements so that it will show the best performance.

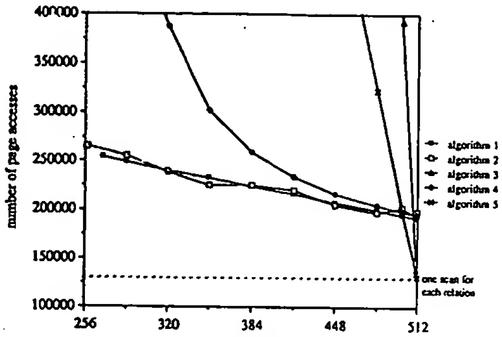


Fig. 12 Simulation Results of Algorithms 1~5 for Non-Uniformly Distributed Relations

4. Extended Join Algorithms

4.1 Description of the Extended Algorithms

From the analysis of the previous five join algorithms, we conclude that, due to insufficient memory space to hold all the necessary pages in each step, the fewer the number of steps, the better the performance. Therefore, algorithms 2 and 4, whose wave propagation speed is higher present better performance. On the other hand, we also verify that the join can be performed with one scan of each relation if the wave propagation speed is low enough. In this case, the join range is too short and the corresponding pages of relation S can be hold in the memory to be used again in the next step. Algorithms 3 and 5 are the only ones to achieve the ideal number of page accesses. We will now modify the join algorithms described above with efforts to :

(a) utilize the memory space efficiently, in order to speed up the propagation of the wave of algorithms 2 and 4 ;

(b) after the wave taken from one relation is loaded in the memory, utilize the remaining memory space efficiently, in order to hold more data of the another relation, for algorithms 1, 3 and 5.

In order to load and unload data in page units, the algorithms presented until now maintain tuples which have already been processed and are not necessary any more in the main memory. We can introduce a garbage collection mechanism which dynamically discards the tuples that are no longer necessary after the processing of each step. This increases the effective space of the main memory for each join step, allowing the loading of more pages in this free space. Following we will analyze how the tuples of relations R and S can be discarded in the previous five algorithms.

First, consider the removal of tuples from relation R. In algorithm 1 and 2, successive waves overlap when propagating over the space of relation R. Thus, a wave of R is loaded in the main memory, used in the processing of one join range and then unloaded, which is then followed by the loading of the next wave. Therefore, for both algorithms the pages of R are discarded as soon as they have been processed in the join range, so that there is no garbage of relation R remaining in the memory. On the other hand, successive waves in algorithm 3 and 4 overlap when propagating over the space of relation R. After processing a join step, it is not necessary to maintain the whole page which overlaps with the successive wave, but only that portion that has not been processed yet.

Now, consider the removal of tuples from relation S. In algorithms 1 and 3, which reserve (M-cs) pages to relation S, there are tuples which have already been processed and are still maintained in the memory. On the other hand, algorithms 2 and 4 reserve only one page for relation S, so that the pages are loaded and then unloaded without garbage maintained in the memory. For algorithm 5, which determines the wave from both relations R and S, it is possible to discard tuples of both relations.

Algorithm	remove tuples of relation R	remove tuples of relation S
1m	X	O
(2m)	X	X
3m	O	O
4m	O	X
5m	O	O

Table 2 Garbage Collection on Extended Algorithms

Excluding algorithm 2, which does not maintain unnecessary tuples in the main memory, we introduce a garbage collection mechanism in the algorithms 1, 3, 4, 5 described above. Here these new algorithms are called algorithms 1m, 3m, 4m and 5m, as shown in Table 2.

The removal of tuples of relation R in algorithm 1m, and of both relations R and S in algorithms 3m and 5m increases the effective space used by relation S. The removal of unnecessary tuples of relation R in algorithm 4m enlarges the effective space used by relation R which means the enlargement of the wave size and the speeding up of the wave propagation. Fig. 13 (a), (b), (c), (d) exemplifies each of these algorithms. After processing the join range in the first step the effective memory space for the next join step is enlarged : the dotted portion is discarded by the garbage collection mechanism and only the dashed portion is maintained in the main memory. In order to clarify this effect, we will present the evaluation of the extended algorithms.

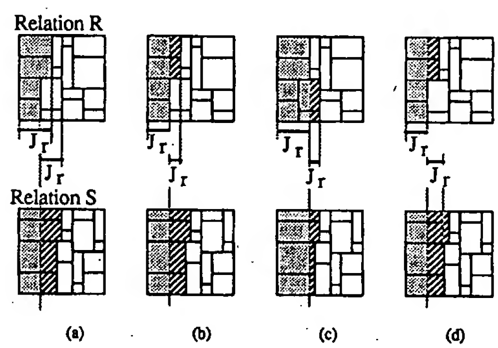


Fig. 13 Garbage Collection on Algorithms (a) 1m (b) 3m (c)4m (d) 5m

4.2 Evaluation for Uniform Data Distribution

4.2.1 Analytical Evaluation

Following, we present the cost formulas for the algorithms presented above. First, we explain the analytical model used for our evaluation. We assumed a simplified model in which :

(1) the KD-tree is perfectly balanced;

(2) the tuples are uniformly distributed within a page; and

(3) for a cluster, denoting μ the mean width of the cluster and σ, the variance, we consider that:

$$\sigma = \frac{(\frac{cs}{2^{k-1}} - 1)\,\delta}{2}, \quad \delta \text{ being a constant.}$$

This means that we consider each of the cs pages of a cluster, except those of the first and last clusters, as having width μ and being dislocated by δ in relation to the previous 2^{k-1} pages in the cluster. This is illustrated by an example for k = 2 in Fig.14 .

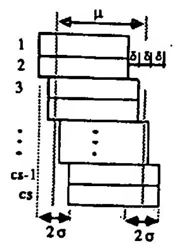


Fig. 14 A Cluster and its mean μ and variation σ

Assuming the model stated above, we can estimate the cost formulas for the extended join algorithms as presented below. For all of them, relation R is read only once, that is, π_R = N pages. To simplify the notation we will use $K = 2^{k-1}$.

Algorithm 1m :

In the same way as algorithm 1, the number of page accesses for relation S is given by expression [2] in section 3.2.1, which we rewrite here :

π_S= (number of join steps) * {(pages within the join range) - (pages remaining in the main memory)}

The part of a cluster of relation S which is processed in a join range and may be maintained in the main memory to be also processed in the next join range, ρ, is represented by the dashed area in Fig. 15 and is calculated as :

$$\rho = \frac{K\delta}{\mu} \sum_{i=1}^{\frac{\alpha}{K}-1} i \text{ pages} = \frac{\alpha}{\mu} cs \text{ pages}$$

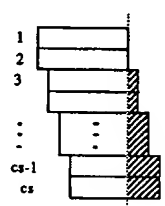


Fig. 15 ρ for Algorithm 1m

Therefore, π_S may be considered for two cases :

(1) ρ fits in the main memory, i.e.,

$$M - cs \geq \frac{\alpha}{\mu} cs$$

In this case, at least the tuples of the dashed area are maintained in the main memory. For this case :

- the number of join steps is equal to the number of clusters, i.e., N/cs ;
- the number of pages of S within the join range is, on average, 2cs ; and
- the number of pages of S which can remain in the main memory is (M - cs - ρ).

Therefore :

$$\pi_S = N\,(3 - \frac{\alpha}{\mu} - \frac{M}{cs})$$

(2) ρ does not completely fit in the main memory, i.e.,

$$M - cs < \frac{\alpha}{\mu} cs$$

For this case :

- the number of join steps is equal to the number of clusters, i.e., N/cs ;
- the number of pages of S within the join range is, on average, 3cs ; and
- the number of pages of S which can remain in the main memory, γ, is calculated by

$$M - cs = \frac{K\delta}{\mu} \sum_{i=1}^{\frac{\gamma}{K}-1} i$$

resulting in :

$$\gamma = \frac{K + \sqrt{K^2 + \frac{4\mu}{\sigma}(M-cs)(cs-K)}}{2}$$

Therefore :

$$\pi_S = N\,(3 - \frac{K + \sqrt{K^2 + \frac{4\mu}{\sigma}(M-cs)(cs-K)}}{2cs})$$

Algorithm 3m :

The join ranges of this algorithm are the same as those of algorithm 3. In this algorithm, unnecessary tuples of both relation R and S are discarded so its cost formula is very complex. Because its expression is too long we omit it here and just show the results in the next subsection.

Algorithm 4m :

In the same way as algorithm 4, here the number of page accesses of relation S is given by expression [3] in section 3.2.1, which we rewrite here :

π_S = (number of join steps) * (pages within the join range)

- the number of join steps is equal to :

$$\frac{N}{\text{new pages in the wave}} = \frac{N}{\alpha}$$

where :

α = (M-1) - (tuples of R remaining in the main memory for the

processing of next join step)

$= (M-1) - \rho$

and where ρ, which is exemplified in Fig.16 as the dashed area, can be expressed as :

$$\rho = \frac{K\delta}{\mu}\left(\sum_{i=1}^{\frac{\alpha}{K}-\frac{x}{K}} i - \sum_{i=1}^{\frac{x}{K}-1} i\right) + x - K$$

where x is :

$$x = \left(\frac{M-1}{cs} - 1\right) cs + K$$

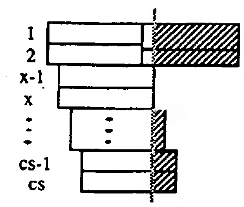


Fig. 16 ρ for Algorithm 3m

Hence, we have that :

$$\rho = \frac{(3\sigma+\mu)cs^2 + (K(\mu-\sigma) + (\mu-2\sigma)(\mu-1))cs - K\mu(\mu-1)}{\mu(cs-K)}$$

and

$$\alpha = \frac{(\mu-3\sigma)cs^2 + (2\sigma(\mu-1) - K(\mu-\sigma))cs}{\mu(cs-K)}$$

- the number of pages of S within the join range is :
  (number of crammed pages of R) + (new pages in the wave)
  $= \chi + \alpha$

where :

$$\chi = (M-1) - \left(\frac{M-1}{cs} - 1\right) cs + K = cs - K$$

Therefore :

$$\pi_S = N\left(\frac{(cs-K)^2\mu}{(\mu-3\sigma)cs^2 + (2\sigma(M-1) - K(\mu-\sigma))cs} + 1\right)$$

4.2.2 Simulation Results

Fig.17 shows the estimation and the respective simulation results of the extended join algorithms presented above, using the same relations R and S described in section 3.2.2. Here again, the prediction was efficient.

Fig.18 shows the simulation results for all the presented algorithms. As expected, all the extended algorithms reduce the number of page accesses in comparison with the basic join algorithms. Algorithm 3m and 5m are the best extended algorithms and they reduce the I/O cost to one scan in most range of variation of the memory size M. Among the basic join algorithms, algorithm 3 and 5 are the only to reduce the I/O cost to one scan, but for the most range of variation of M, they show the highest I/O cost because they have the lowest wave propagation speed.

The introduction of the garbage collection mechanism increases the effective memory space. For the basic join algorithms, when the memory space was not enough to save the necessary pages in each join step, the performance degraded as the number of join steps increased. Algorithms 1m, 3m and 5m increases the effective memory space to save more tuples of relation S while algorithm 4m enlarges the wave. For the extended join algorithms, the saving of tuples of S are more efficient than the decreasing of the number of join steps. The wave propagation speed increases from 5m, 3m, 1m to 4m and the I/O cost increases in this sequence.

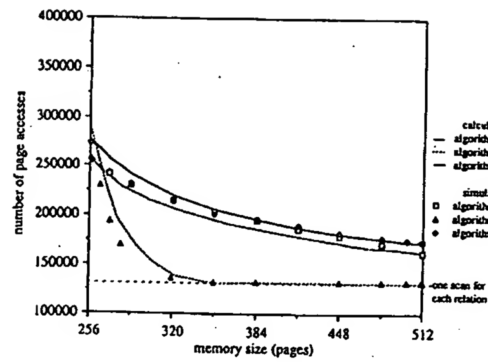


Fig.17 Analytical and Simulation Results of Algorithms 1m~4m for Uniformly Distributed Relations

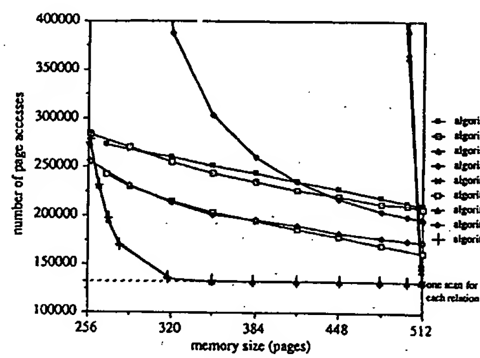


Fig.18 Simulation Results of Algorithms 1~5, 1m~5m for Uniformly Distributed Relations

4.3 Simulation Results for Non-Uniform Data Distribution

In this simulation we used the same relations described in section 3.4.1 and the results are shown in Fig.19. In the same way as for the uniform data distribution, the I/O cost increases for algorithms which reserve more memory space to be used by the wave. For the case of normal and inverse normal data distribution, the fact of choosing the appropriate wave from the KD-tree information of both relations is well utilized : algorithm 5m resulted the best in the entire range of main memory size, minimizing the I/O cost to one scan of each relation.

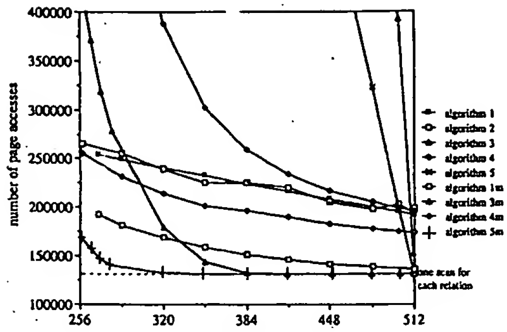


Fig. 19 Simulation Results of Algorithms 1~5, 1m~5m for Non-Uniformly Distributed Relations

5. Conclusion

In this paper we investigated join algorithms for KD-tree indexed relations, presenting their analytical and simulation results. KD-tree is a multidimensional clustering mechanism based on the adaptable method and so different clusters have tuples with a same attribute value. The join of two relations R and S can not be simply reduced to the join of one of their respective cluster and a naive join algorithm for KD-tree indexed relations requires a high I/O cost. An efficient join algorithm for KD-tree indexed relations is complex and has not been investigated until now. In this paper we propose efficient join strategies to reduce its I/O cost and extensively evaluate these algorithms with analytical analysis and simulation. With one of our proposed join algorithms which uses a suitable memory management, the join of two non-resident relations can be performed with one scan.

According the the KD-tree information, the concepts of wave and join range are introduced and, based on them, the join algorithms are proposed. Four basic algorithms determine the wave and join ranges according to the KD-tree information of relation R and another algorithm determines them with the information of both relations R and S. Analytical formulas and simulation results clarify their characteristics. For these basic join algorithms, the higher the wave propagation speed, the better the algorithm performance. This is because in each step, the memory space to maintain the necessary tuples for the next step is small and insufficient, and so the fewer the number of steps, the fewer the number of scans of relation S. Therefore, we propose extended versions of the basic algorithms, introducing a garbage collection mechanism to enlarge the effective memory size. For these algorithms, the tuples that have already been processed and are no longer necessary are discarded from the main memory at the end of each join step. These extended algorithms are analyzed by analytical formulas and simulation results. In opposition to the former basic algorithms, in these extended algorithms, as the wave propagation speed is increased, the space reserved for relation S decreases and the performance worsens. For the best algorithm, the unnecessary tuples of both relations are dynamically discarded to enlarge the effective memory space and the wave is determined so that its speed of propagation is high. With this algorithm, the join can be performed with almost one scan of each relation.

The presented figures were only for KD-trees of two dimensions but although not exposed here, analysis of higher dimensions were performed, also showing that the analytical estimation matches the simulation results.

The implementation details of these join strategies on KD-tree indexed relations are to be reported in a future paper.

References

[1] J.L.Bentley : "Multidimensional Binary Search Trees Used for Associative Searching", Commun. ACM, 18, 9, pp.509-517 (September 1975).

[2] J.L.Bentley : "Multidimensional Binary Search Trees in Database Applications", IEEE Trans. Software Eng., 5, 4, pp.333-340 (1979).

[3] J.M.Chang and K.S.Fu : "A Dynamical Clustering Technique for Physical Database Design", Proc.of the 1980 SIGMOD Conf., pp.188-199 (1980).

[4] J.P.Cheiney, P.Faudemay, R.Michel and J.M. Thevenin : "A Reliable Parallel Backend Using Multiattribute Clustering and Select-Join Operator", Proc. of 12th. Int. VLDB Conf., pp.220-227 (August 1986).

[5] S.Fushimi, M.Kitsuregawa, M.Nakayama, H.Tanaka and T.Moto-oka : "Algorithm and Performance Evaluation of Adaptive Multidimensional Clustering Technique", Proc. of the 1985 SIGMOD Conf., pp.308-318 (1985).

[6] A.Guttman : "R-Trees : A Dynamic Index Structure for Spatial Searching", Proc.of the 1984 SIGMOD Conf., pp.47-57 (June 1984).

[7] K.L.Kelley and M.Rusinkiewicz : "Implementation of Multi-Key Extendible Hashing as an Access Method for a Relational DBMS", Proc.of 2nd. Data Engineering Conf., pp.124-131 (February 1986).

[8] H.P.Kriegel and B.Seeger : "Multidimensional Dynamic Quantile Hashing is Very Efficient for Non-Uniform Record Distributions", Proc.of 3rd. Data Engineering Conf., pp.10-17 (February 1987).

[9] H.P Kriegel and B.Seeger : "PLOP-Hashing : A Grid File without Directory", Proc.of 4th. Data Engineering Conf., pp.369-376 (February 1988).

[10] J.Nievergelt, H.Hinterberger and K.C.Seveik : "The Grid-File : An Adaptable, Symmetric Multikey File Structure", ACM Trans. Database Syst., 9, 1, pp.38-71 (March 1984).

[11] E.J.Otoo : "A Mapping Function for the Directory of a Multidimensional Extendible Hashing", Proc. of 10th. Int. VLDB Conf., pp.493-506 (August 1984).

[12] E.J.Otoo : "A Multidimensional Digital Hashing Scheme for Files With Composite Keys", Proc.of the 1985 SIGMOD Conf., pp.214-229 (June 1985).

[13] E.A.Ozkarahan and M. Ouksel : "Dynamic Order Preserving Data Partitioning for Database Machines", Proc. of 11th. Int. VLDB Conf., pp.358-368 (August 1985).

[14] E.A.Ozkarahan and H. Bozsahin : "Join Strategies Using Data Space Partitioning", New Generation Computing, 6, 1, pp.19-39 (1988).

[15] J.T.Robinson : "The KDB-Tree : A Search Structure for Large Multidimensional Dynamic Indexes" Proc.of the 1981 SIGMOD Conf., pp.10-18 (1981).

[16] Y.Tanaka : "Massive Parallel Architecture for Very Large Databases", Doctoral Thesis, Univ. of Tokyo (May 1985).

[17] J.A.Thom, K.Ramamohanarao and L. Naish : "A Superjoin Algorithm for Deductive Databases", Proc.of 12th. Int. VLDB Conf., pp.189-196 (August 1986).

[18] P.Valduriez and Y. Viemont : "A Multikey Hashing Scheme using Predicate Trees", Proc.of the 1984 SIGMOD Conf., pp.107-114 (June 1984).

Join strategies on KD-tree indexed relations

Kitsuregawa, M. Harada, L. Takagi, M.
Inst. of Ind. Sci., Tokyo Univ.;

This paper appears in: **Data Engineering, 1989. Proceedings. Fifth International Conference on**
Publication Date: 6-10 Feb 1989
On page(s): 85-93
Meeting Date: 02/06/1989 - 02/10/1989
Location: Los Angeles, CA, USA
ISBN: 0-8186-1915-5
References Cited: 18
INSPEC Accession Number: 3406286
DOI: 10.1109/ICDE.1989.47203
Posted online: 2002-08-06 16:46:03.0

Abstract
Join algorithms on KD-tree indexed relations are proposed. The join algorithms are based on a concept called wave. Th wave is a set of pages that is the object of joining and that propagates over the relation space in the direction of the join attribute axis. Four basic join algorithms that determine the wave from one of the relations and one algorithm that determines the wave from both relations are proposed. The algorithms are described and extensively analyzed with analytical formulas and simulation results. Then a garbage collection mechanism is introduced that discards the unnecessary data loaded in the main memory and extends the previous basic algorithms with an efficient memory management. It is shown that the proposed algorithms perform the join of very large relations with one scan

Index Terms
Inspec
Controlled Indexing
relational databases storage management

Non-controlled Indexing
KD-tree indexed relations garbage collection mechanism join algorithms join attribute memory management query processing relation space wave

Author Keywords
Not Available

References

No references available on IEEE Xplore.

Citing Documents

1 Utilizing page-level join index for optimization in parallel join execution, Chiang Lee; Zue-An Chang
Knowledge and Data Engineering, IEEE Transactions on
On page(s): 900-914, Volume: 7, Issue: 6, Dec 1995

09/146,002

1 23
2 3 4 5 6 7 8 9 10
2 3 4 5 6 7 8 9 10
6 7

11 24

18 25
13 14 15
12 13 15

16 26
17 17

18 27
19 28
18 20

21 29
1
20 22

Draft of claims. Please do not Remove until at Publications! Thank you.